Sede Amministrativa: Università degli Studi di Padova
Dipartimento di Fisica

DOTTORATO DI RICERCA IN FISICA

CICLO XII

Performance of the new tracking system at CDF II

Coordinatore: Prof. Attilio Stella

Supervisore: Prof. Dario Bisello

Dottorando: Paolo Gatti

31 Dicembre 2000

Introduction

Tracking down charged particles, as they traverse the numerous layers of a modern collider detector, is an important part of event analysis. When a collision causes the production of hundreds of particles, as can be the case in hadronic colliders, the task of associating the correct points to each track becomes particularly difficult and time-consuming.

The subject of this thesis is the reconstruction of tracks at the CDF II experiment, which will soon begin collecting data on $p\bar{p}$ collisions at the Tevatron accelerator, at a center-of-mass energy of 2 TeV. The first phase of the CDF experiment ended in 1996; during the five-year shutdown, the accelerator and detectors were thoroughly upgraded.

I begun research on CDF tracking when I was preparing my laurea thesis (1996), and I continued to pursue this subject during my graduate studies. I devoted most of my time to the "Histogram Tracking" algorithm, which will be described in detail in Chapters 7 and 8. At first I performed some feasibility tests, in order to select the most promising approaches; starting in 1999, I implemented Histogram Tracking within the official CDF II reconstruction code, and I compared its performance to other algorithms.

A second kind of activity consisted in the development of testing tools for the CDF II simulation, in the optimization of some particularly time-consuming parts of the code, and in the development of "utility" C++ classes. Although these issues are not discussed in the thesis, the related work was quite useful in detecting and correcting code bugs, and in ensuring the simulation would run smoothly.

Lastly, I examined the performance of SVT, the hardware tracker which is used in the CDF II trigger to provide early detection of secondary vertices. At first, I made sure the SVT simulation would give reasonable results when applied to the baseline detector geometry. Then, I studied the consequences of using Layer 00 (a recently added silicon layer, placed just outside the beam pipe) as an input to SVT.

The thesis is structured as follows.

Part I describes the CDF II experiment and some of its goals in the b physics sector.

Chapter 1 is an introduction to the current status of b physics at hadronic colliders: how B hadrons are produced, how they decay, how their properties are measured, and why an efficient reconstruction of tracks is particularly relevant to b analyses. In order to improve the current results, the detector must satisfy certain requirements, which I state at the end of the chapter, and which were used as design goals for CDF II.

Chapter 2 describes the upgraded structure of the Tevatron and of CDF II from a global point of view; in **Chapter 3** I focus on the central tracking systems, describing them in detail.

Part II is devoted to the use of tracking in the experiment trigger. In **Chapter 4** I briefly describe XFT, the fast trigger operating on the drift chamber; in **Chapter 5** I introduce SVT, the Silicon Vertex Tracker, and present the results of my studies on its performance.

Part III, lastly, deals with the use of the tracking system in the final level of trigger and in offline analyses.

In **Chapter 6** I describe the tools I used to determine the tracker's performance: Monte Carlo samples, a "best case" performance analysis, and some conventions on how to classify hits and tracks for the purpose of deciding whether a track "has" been found and "should have" been found.

Chapter 7 deals with the algorithms used in reconstructing tracks inside the COT, CDF II's drift chamber. I present the "Segment Linking" strategy, which was used as baseline in the planning of CDF II, and the "Histogram Tracking" approach, whose implementation and testing represented a large part of my research. I then compare the two strategies in terms of efficiency and time consumption.

Finally, **Chapter 8** describes the tracking algorithms used in the silicon vertex detector. I begin by outlining the reconstruction of clusters, and their use in determining the primary vertex position. After an introduction to the silicon tracking framework, I present the three main algorithms used in track reconstruction: Stand-alone, Outside-In, and Histogram Tracking. In the end, I examine the performance of silicon tracking at CDF II.

Contents

III   Full Track Reconstruction   89

6   Data samples and analysis tools   91

7   Tracking in the COT   109

8   Tracking in the silicon   137

Part I

The CDF II Experiment

Chapter 1

B Physics at CDF II

The bottom quark was unexpectedly discovered in 1977 at Fermilab [1], where the Υ resonance was first observed in a $\mu^+\mu^-$ mass spectrum. Its detection opened the way to a new sector of particle physics: it provided evidence for a third generation of quarks and leptons in the Standard Model, and led, for example, to a theoretical basis to explain CP violation.

One peculiarity of the b quark is its very weak coupling to the lighter, lower-generation quarks: the charm-bottom transition amplitude is CKM—suppressed by a factor of about 0.04 with respect to u—d or c—s amplitudes. As a consequence, B hadrons are "long-lived", with average lifetimes of the order of 1.5 ps; a relativistic B hadron can move several hundred microns before decaying in an observable secondary vertex.

B mesons and baryons are therefore an important probe for the study of weak interactions: they are the best source of information for the five CKM matrix elements involving third-generation quarks. Such hadrons have been studied extensively in the past two decades, both in general-purpose experiments and at b factories. They still represent part of the frontier of particle physics: their study is therefore a central part of the physics program of new experiments, such as CDF II.

1.1   B physics at hadronic colliders

The easiest way to obtain a clean B sample consists in using an e^+e^- collider, either at the $\Upsilon(4S)$ peak (as in BaBar/Belle) or at the Z^0 pole (LEP/SLC). In these conditions, $b\bar{b}$ pairs are produced in the s channel only, with a cross section of about 1 nb and 6 nb respectively.

At a hadronic collider, on the other hand [2], b production can also take place (at tree level) via gluon–gluon fusion; higher order processes, such as

flavour excitation and gluon splitting, are also relevant. The b production cross section can be quite high: at the Tevatron, at $\sqrt{s} = 1.8$ TeV, the p_t-integrated production cross section in the rapidity [1] region $|\eta| < 1$ is of the order of 50 μb [3].

Another advantage of $p\bar{p}$ colliders is the higher center-of-mass energy. This contributes to the increase of the production cross section; it also allows the study of heavy B hadrons, such as Λ_b, whose production is kinematically forbidden at the $\Upsilon(4S)$ resonance.

The main drawback of hadronic colliders is their large total cross section for inelastic $p\bar{p}$ scattering, which exceeds the $b\bar{b}$ production cross section by three orders of magnitude. In order to take advantage of the increased b production rate, it is necessary to improve the signal to noise ratio. The detector must therefore provide sufficient information to perform this task; the event reconstruction and trigger systems must take advantage of this information.

1.1.1   Production of b quarks

At the center-of-mass energy of the Tevatron, $p\bar{p}$ collisions can be seen as collisions between two beams of almost free partons: valence (anti)quarks, gluons, sea quarks and antiquarks. Any component of the p beam can interact with any component of the $\bar{p}$ beam, giving rise to a high number of Feynman diagrams.

At the leading order in α_s, as shown in figure 1.1, $b\bar{b}$ pairs are produced by $q\bar{q}$ annihilation (s channel, color singlet) or, more frequently, by gluon fusion (in any of the s, t, u channels; both as a color singlet and as an octet).

Incoming partons have a negligible transverse momentum, and there are only two outgoing particles; as a consequence, the b and $\bar{b}$ quarks are back to back in the transverse plane[2]. The rapidity difference between the two quarks is usually of the order of one or less; as a consequence, a large fraction of events has at least one b quark in the central region ($|\eta| < 1$) of the detector.

The average transverse momentum of the b quarks thus produced turns out to be of the same order of magnitude of the b mass, that is about 5 GeV; the p_t spectrum has a steep descent above the peak.

---

[1]By definition, the rapidity of a particle is $y \equiv \frac{1}{2}\log\left(\frac{E+p_z}{E-p_z}\right)$, where E is the energy and p_z the longitudinal momentum in the detector's reference frame. For ultrarelativistic particles, y can be approximated with the pseudorapidity $\eta \equiv -\log\tan(\theta/2)$, θ being the polar angle.

[2]Since the incoming partons can have different longitudinal momenta, the $b\bar{b}$ system is usually boosted along the z axis.
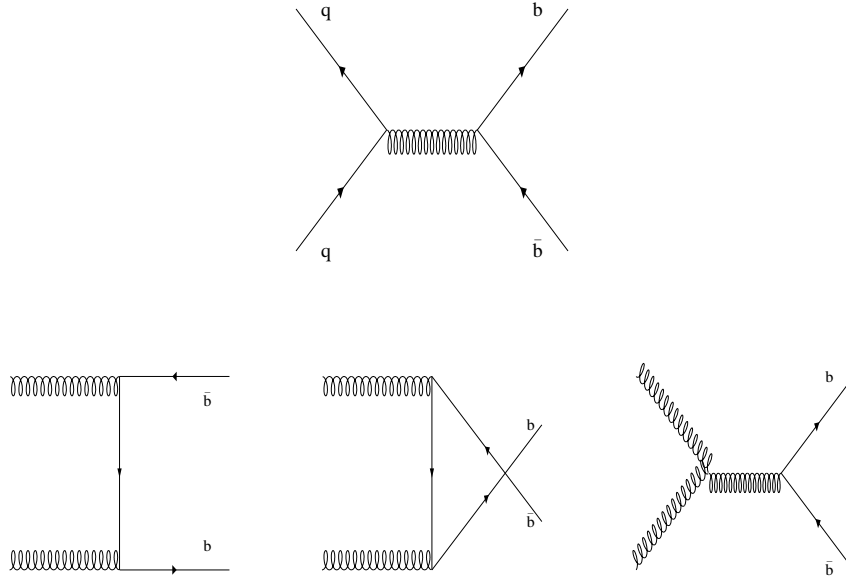
Figure 1.1: Tree-level production of $b\bar{b}$ pairs, by $q\bar{q}$ annihilation (above) or gluon fusion (below).

Soon after the discovery of the b quark [4], next-to-leading order contributions to $b\bar{b}$ production were computed. They were found to be far from negligible: at high energies, gluon—gluon interaction terms can be orders of magnitude larger than those associated to $q\bar{q}g$ vertices. When $\sqrt{s} \gg m_b$, as is the case at CDF II, NLO terms are actually larger than LO contributions.

Some of the NLO processes involve the emission of a real gluon from one of the leading order diagrams; during hadronization, this leads to the formation of three jets. In other cases, a virtual gluon is emitted and reabsorbed; these diagrams interfere with the corresponding leading-order process.

It is also possible, as shown in figure 1.2, to have NLO diagrams with a quite different kinematic structure. After a gg or qg interaction, a gluon can split in a $q\bar{q}$ pair. If this is a $b\bar{b}$ pair, the two quarks usually have similar directions and energies; the resulting jets often merge into a single jet.

Another possibility consists in flavour excitation, also shown in the figure. In this case the $b\bar{b}$ pair is produced by fusion of a real gluon with a virtual gluon, in the t or u channel. As a result, the two b quarks can have a large difference in transverse momentum.
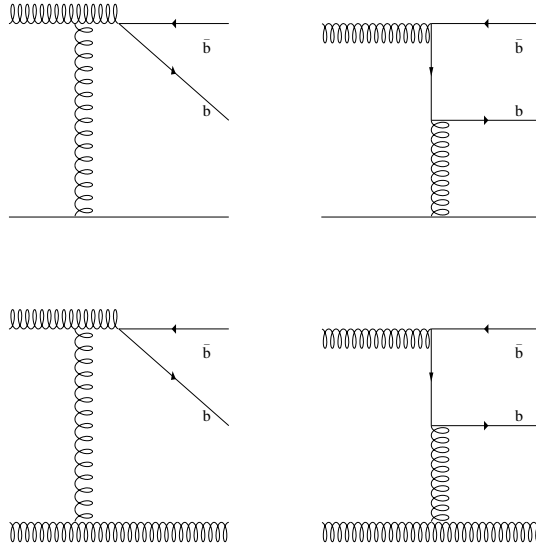
Figure 1.2: Some next–to–leading–order production diagrams of $b\bar{b}$ pairs: gluon splitting (left) or flavour excitation (right).

1.1.2   Parton densities

By applying the Feynman rules to the diagrams of the previous section, one obtains amplitudes and cross sections for the elementary, parton-level processes. The results of these computations must be related to the observable cross sections, which involve collisions between protons and antiprotons — that is, between composite objects. This problem is solved by resorting to **parton distribution functions** (PDF).

When a hadron a (for example a proton) of given momentum interacts with a probe, the probe actually scatters against a parton of kind i (a gluon, or any flavour of quark or antiquark) which carries a fraction x of the hadron momentum. The parton distribution function $F_i^a(x, Q^2)$ describes the probability of this process, as a function of x and of the total momentum transfer Q^2. At high Q^2, the dependence of F on Q^2 becomes small. This behaviour is called Bjorken scaling; it provides experimental evidence for the existence of point-like constituents in baryons.

Parton distribution functions are measured [5] by analyses of deep inelastic scattering data (provided by experiments such as NMC [6], H1 [7] and ZEUS [8]) and of the inclusive jet cross section in hadronic collisions [9]. Among other things, these measurements point out that about half of

the momentum of an (anti)proton is carried by gluons, almost irrespectively of the Q^2 scale. At low energies, the remaining momentum belongs almost completely to the valence quarks; as energy increases, sea (anti)quarks and heavy quarks become more important.

In order to obtain the overall cross section for the production of b quarks in $p\bar{p}$ collisions, it is necessary to weigh the parton-level cross section with the appropriate PDF, integrate them over the incoming parton momentum fractions and the outgoing quark momentum, and sum over all the possible parton combinations:

$$\frac{E\,d^3\sigma}{dp^3} = \sum_{ij} \int dx_i dx_j \left(\frac{E d^3 \hat{\sigma}_{ij}}{dp^3}(x_i, x_j, p, \mu) \right) F_i^a(x_i, \mu^2) F_j^b(x_i, \mu^2) \qquad (1.1)$$

The actual energy scale of the process, μ, is unknown; it is a major source of systematic error on QCD predictions of heavy quark production.

1.1.3 Hadronization

Once a pair of energetic heavy quarks has formed, as described above, it undergoes a nonperturbative hadronization process. The original quarks are coloured objects, bound together by the strong interaction; in order for them to move apart, they must bind to other quarks and form colour singlets.
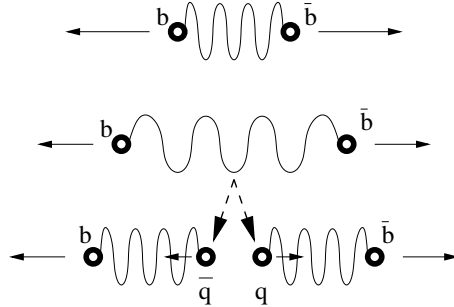
Figure 1.3: String fragmentation model

This situation can be schematized by a string fragmentation model [10], in which the attraction between b and $\bar{b}$ is mediated by a string of virtual gluons. When the quark-antiquark distance reaches a certain threshold, the string breaks, and part of its energy is transformed into a new $q\bar{q}$ pair, as schematized in figure 1.3. In turn, the two new strings stretch and fragment; the process continues as long as there is enough energy for quark pair creation.

Unless fragmentation stops immediately after the first step, part of the original b quark momentum is transferred to the intermediate mesons. The probabilistic relationship between b quark and B hadron momenta is described by fragmentation functions, as the ones studied in [11].

The situation is more complicated when the b and $\bar{b}$ quarks do not have opposite colour, as happens for instance when they are produced by gluon splitting, or when a B baryon (instead of a meson) is produced in the final state. In these cases, in order to obtain colour singlets in the final state, it is necessary for a b quark to exchange a gluon with some other component of the beam.

As a result of fragmentation, the original b and $\bar{b}$ quarks can form several kinds of hadrons, the lightest of which are the pseudoscalar mesons $B^0 = \left| \bar{b}d \right\rangle$, $B^+ = \left| \bar{b}u \right\rangle$, $B_s^0 = \left| \bar{b}s \right\rangle$ and $B_c^+ = \left| \bar{b}c \right\rangle$, the $\Lambda_b = |bdu\rangle$ baryon, and their charge conjugates.

1.1.4 B decay

Bottom quarks decay, by emitting a virtual W, into charm or up quarks. These processes are CKM–suppressed, while decay into a top quark is kinematically forbidden. The b quark has therefore a "long" lifetime of the order of 1.5 ps, which allows a relativistic B hadron to travel an average of several hundred μm before decaying in a secondary vertex.

Due to the high mass difference between bottom and charm quarks, B hadron decay can occur via a large number of exclusive channels. For this reason, many b-physics analyses study either an inclusive sample, or a few specific exclusive channels whose characteristics make event reconstruction particularly clean.

In most cases, the W boson decays into a $q\bar{q}'$ pair. This can happen in several ways, as shown in figure 1.4; some of these diagrams, like internal and external W emission, interfere with each other. Fragmentation can also occur during W decay, leading to the production of one or more additional light mesons.

About 10% of B mesons decay semileptonically; the final state comprises a lepton, the corresponding neutrino, and one or more hadrons — one of which is most frequently a D meson. These events are easier to distinguish from the background, due to the lepton and the missing energy; they are also less affected by theoretical uncertainties, since the W decay products do not interfere with the hadronic branch of the diagram.

Pure leptonic decay modes are also possible for charged B mesons, via annihilation into a W. These channels are helicity-suppressed; a 90% CL
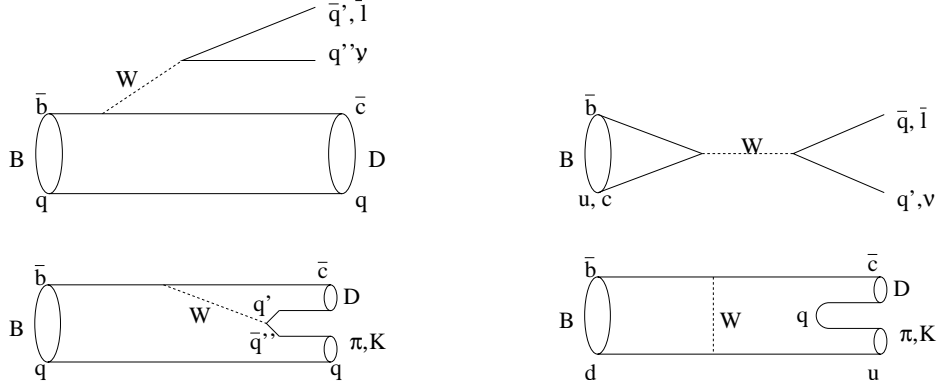
Figure 1.4: Some B meson decay modes: external (top left) and internal (bottom left) W emission, B^+ annihilation (top right), W exchange (bottom right).

limit to their branching ratio [12] is $< 1.5\,10^{-5}$ for $B^+ \to e\nu_e$ and $< 2.1\,10^{-5}$ for $B^+ \to \mu\nu_\mu$.

To end with, a particularly relevant family of hadronic exclusive channels consists of B decaying into a J/ψ and a light meson, such as a K, π, ρ or η. When the J/ψ decays to a $\mu^+\mu^-$ pair, the event signature is especially clear, and can be used to select the event at trigger level.

1.2 B triggering and identification

B identification is a critical step, especially in a hadronic collider environment, where the $b\bar{b}$ production cross section is only $\sim 10^{-3}$ of the total inelastic cross section. It is essential to increase the signal to background ratio.

In order to collect sufficient statistics on rare processes in a reasonable time, modern colliders operate at a high frequency; since storage space and data transfer rates are limited, the detector must be able to trigger quickly on the few interesting events, while discarding a large fraction of the background.

One effective technique consists in the use of multi-level triggers. A very fast and rough algorithm, usually implemented in hardware, examines the response of the fastest parts of the detector; it provides a first improvement of the S/B ratio and eliminates the vast majority of uninteresting events, making it possible to perform more elaborate and time-consuming analyses on the surviving events.

1.2.1   Single lepton trigger

The simplest trigger for b physics consists in the search of a charged lepton from a semileptonic B decay.

Electrons are identified by their release of energy in the electromagnetic compartment of the calorimeter[3], in coincidence with a charged track pointing at the cluster. At a higher trigger level, the shower profile in the calorimeter can also be taken in consideration; electrons from γ conversions can be filtered out, by looking for pairs of opposite-charged tracks separated by a small angle.

Muons are instead detected by their penetrating power. When they cross detector material, their energy release is compatible with minimum ionization; their trajectory can be tracked both inside and outside the calorimeter, provided that a certain amount of multiple scattering is accounted for.

Since the p_t spectrum of b quarks decreases sharply at energies above m_b, an efficient single-lepton trigger must have a low p_t threshold; this is not always possible, if the fake rate is to be kept small.

1.2.2   Dilepton trigger

A second kind of b trigger requires the simultaneous presence of two leptons. This requirement strongly decreases the background, allowing the use of much lower energy and momentum thresholds; it also entails a smaller geometric acceptance region, as both leptons from the event need to lie inside the detector's active region. A wide tracking and muon coverage is therefore particularly advantageous for this kind of trigger.

Dimuon triggers are simple, fast, efficient, and have a low background rate. Their first requisite is the presence of two opposite-charged tracks, each matching a stub in the muon chambers outside the calorimeter. It is then possible to compute the invariant mass of the $\mu^+\mu^-$ system, and use it to classify the event as containing a J/ψ, a $\psi(2S)$, a Υ, or none of the above.

Dielectron and muon-electron triggers are also feasible. Since electrons have a higher background than muons, these triggers require a higher p_t threshold and provide a smaller sample.

1.2.3   Secondary vertex trigger

As was stated above, relativistic B hadrons traverse an average distance of several hundred microns before decaying. If the detector is capable of providing accurate track position measurements close to the interaction point,

---

[3]The CDF II calorimeter is described in section 2.5.

it is possible to recognize B events by looking for a secondary vertex. This can be done by projecting the event on the transverse plane, or in a full 3D reconstruction; the latter option requires a more complex detector and takes longer.
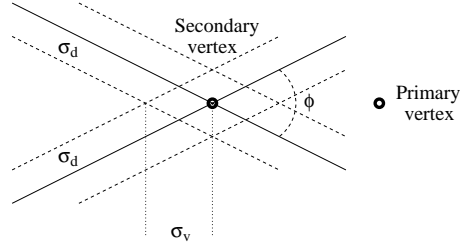
Figure 1.5: Secondary vertex resolution on the transverse plane

The secondary vertex position resolution depends strongly on the number of reconstructed tracks starting in that vertex, on the angles they form among each other, and on the single track resolution. Figure 1.5 exemplifies a symmetric, two-track vertex on the transverse plane. Each track's position is known with a resolution σ_d; the vertex-to-vertex distance on the transverse plane has an error of about $\sigma_d/\sin(\phi/2)$, where ϕ is the angle between the track extrapolations. Vertex resolution obeys a more complex formula in the general case; it remains true, however, that the best results are obtained when secondary tracks form wide angles among each other.

The mass difference between a B hadron and its decay products is usually large, of the order of 3 GeV; the p_t spectrum of B hadrons decreases sharply at energies above 5 GeV. As a consequence, it is possible for the B decay products to form a rather wide angle; the secondary vertex position can be measured with good accuracy by intersecting two tracks.

If the number of tracks in the event is large, examining all the secondary vertex candidates can become a lengthy process, which exceeds the time slice allotted to a low-level trigger. It is possible, however, to trigger on the simple presence of one or two tracks with a large displacement from the primary vertex: such tracks have necessarily origin in a displaced secondary vertex. The opposite is not true: if the tracks from a displaced secondary vertex form small angles with respect to each other, they can all pass quite close to the primary vertex as well.

The study of certain phenomena, such as CP violation and oscillations of neutral B, benefits from detection of secondary vertices at very little displacements from the primary vertex. On the other hand, vertex resolution is finite; the track displacement threshold must be large enough to discard a

substantial fraction of background events.

A major vulnerability of secondary vertex triggers is their sensitivity to occupancy: high occupancy on the innermost detectors can lead to the reconstruction of fake tracks, whose displacement from the primary vertex can be quite large. The minimum required number of displaced tracks and the displacement threshold must therefore be adapted to the expected occupancy.

Once the event has passed the lower trigger levels, it becomes possible to perform a more complete reconstruction, determining which tracks originate in each vertex. If the particle momenta and energies are known, they can be used to determine the invariant mass of the decayed particle, and to assign the event to a specific, exclusive channel.

1.3   B hadron lifetimes

In the simplest model of B hadron decay, the b quark decays as if it were a free quark; the other quarks in the original hadron act as spectators. If this model were accurate, B meson and baryon average lifetimes would obey the equation

$$\tau = \frac{192\,\pi^3}{G_F^2\,m_b^5\,|V_{cb}|^2\,\mathcal{F}}$$ (1.2)

where $\mathcal{F}$ is a phase space factor. This is not the case: the interaction between the constituent quarks of a hadron is not negligible. Some of the B decay modes shown in figure 1.4 involve the non-b quark directly; some other diagrams interfere with each other.

The first measurement of B lifetimes [13] [14] were performed by extrapolating the B decay products back towards the primary vertex, as in figure 1.6, and estimating the impact parameter d.

For a given value of ϕ (the angle between the B momentum and the secondary track), d is proportional to the path L the B hadron traversed before decaying; this, in turn, is equal to the proper lifetime τ multiplied by $c\beta\gamma$. On the other hand, $\tan\phi$ is inversely proportional to the total momentum along the B direction, $p_B = m_B c\beta\gamma$. As a consequence, the impact parameter

$$d = L\sin\phi \simeq L\tan\phi = (\tau c\beta\gamma)\frac{\alpha}{c\beta\gamma} = \alpha\tau$$ (1.3)

is proportional to the B lifetime; its dependence on the B momentum is weak, being proportional to $\cos\phi$. It is therefore possible to extract τ from a fit to the observed d distribution.
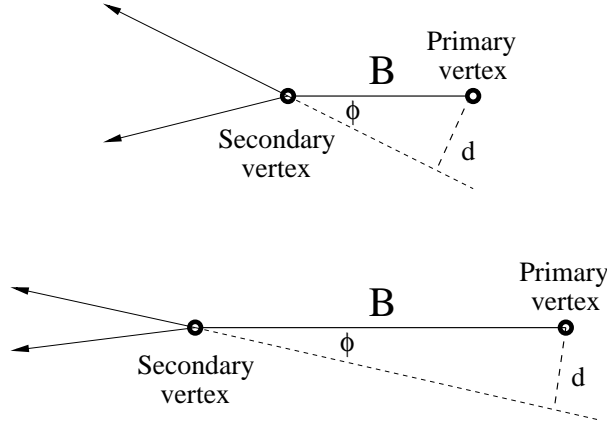
Figure 1.6: Decay of two B hadrons with $\beta\gamma$ equal to 1 (above) or 2 (below). With higher boost, the B hadron lives longer, but its decay products form a smaller angle; the impact parameter d is almost the same.

An advantage of this procedure consists in its simple requirements: it only needs the primary vertex location and a single reconstructed track. As a consequence, it can be applied directly to partially reconstructed events — such as semileptonic decays, where the neutrino is not observable.

A more complex method requires an independent measurement of L and $c\beta\gamma$; their ratio provides the proper B lifetime. To determine L, it is necessary to reconstruct the event geometry. The secondary vertex position is determined by intersecting the backwards extrapolations of tracks. The primary vertex can be measured in the same way; as an alternative, its position on the transverse plane can be approximated with the center of the beam spot. The former choice would be better in case of a perfectly reconstructed event; however, the primary vertex fit might be contaminated by the decay products of the *other* b quark. If the beam spot is small, with a radial σ of the order of 25 μm, the latter choice provides a good approximation without introducing a possible bias.

As for $c\beta\gamma$, it is equal to the ratio between the B momentum and its mass. This can be measured directly only in fully reconstructed events, where all the decay products are visible; in semileptonic decays it is necessary to infer the neutrino momentum by a Monte Carlo simulation, as a function of visible momenta.

Published analyses have made use of both methods, achieving similar accuracies; the current world averages, as presented by the LEP B lifetimes working group [15], are shown in table 1.1. The same reference provides a

listing of analyses on specific B hadrons — B^0, B^+, B_s^0 and Λ_b. Preliminary
B_c results are given in [16].

Experiment	Method	Dataset	τ_{B} (ps)	Ref
ALEPH	Lepton i.p. (3D)	91-93	$1.533 \pm 0.013 \pm 0.022$	[17]
L3	Lepton i.p. (2D)	91-94	$1.544 \pm 0.016 \pm 0.021$	[18][b]
OPAL	Lepton i.p. (2D)	90-91	$1.523 \pm 0.034 \pm 0.038$	[19]
Aver. set 1			1.537 ± 0.020	
ALEPH	Dipole	91	$1.511 \pm 0.022 \pm 0.078$	[20]
DELPHI	All track i.p.(2D)	91-92	$1.542 \pm 0.021 \pm 0.045$	[21][a]
DELPHI	Sec. vert.	91-93	$1.582 \pm 0.011 \pm 0.027$	[22][a]
L3	Sec. vert. + i.p.	91-94	$1.556 \pm 0.010 \pm 0.017$	[18][b]
OPAL	Sec. vert.	91-94	$1.611 \pm 0.010 \pm 0.027$	[23]
SLD	Sec. vert.	93	$1.564 \pm 0.030 \pm 0.036$	[24]
Aver. set 2			1.577 ± 0.016	
CDF	J/ψ vert.	92-95	$1.533 \pm 0.015\,^{+0.035}_{-0.031}$	[25]
Average			1.564 ± 0.014	

Table 1.1: Measurements of the average B hadron lifetime, as quoted by [15].
a) The combined DELPHI result quoted in [22] is $1.575 \pm 0.010 \pm 0.026$ ps.
b) The combined L3 result quoted in [18] is $1.549 \pm 0.009 \pm 0.015$ ps.
The analyses have been divided in three groups, which, due to b production
or to sample selection issues, differ in the relative amount of the various B
hadron types.

1.4 B^0 oscillations and mixing

1.4.1   Generalities
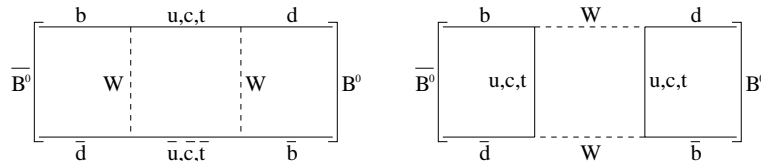
Figure 1.7: Neutral B meson oscillations

Neutral mesons, such as K^0 [26] and B^0, can oscillate into their own antiparticles via a double W exchange, as in the box diagrams of figure 1.7. The mass eigenstates of the neutral B system differ from the flavour eigenstates: if CP violation is neglected, the mass eigenstates are

$$|B_1\rangle \equiv \frac{1}{\sqrt{2}}\left(\left|B^0\right\rangle - \left|\bar{B}_0\right\rangle\right) \tag{1.4}$$

$$|B_2\rangle \equiv \frac{1}{\sqrt{2}}\left(\left|B^0\right\rangle + \left|\bar{B}_0\right\rangle\right) \tag{1.5}$$

$$\tag{1.6}$$

Due to the mass difference Δm between B_1 and B_2, the corresponding wavefunctions oscillate with different frequencies, leading to a phase difference increasing linearly with time. Therefore, starting with a pure B^0 state at $t = 0$, the probabilities of observing a B^0 or a $\bar{B}^0$ at proper time t is, as plotted in figure 1.8,

$$\begin{aligned}
\mathcal{P}_{mix} \equiv \mathcal{P}(B^0 \to \bar{B}^0, t) &= \frac{1}{2}e^{-\Gamma t}(1 - \cos \Delta m\, t) \\
\mathcal{P}_{unmix} \equiv \mathcal{P}(B^0 \to B^0, t) &= \frac{1}{2}e^{-\Gamma t}(1 + \cos \Delta m\, t)
\end{aligned} \tag{1.7}$$
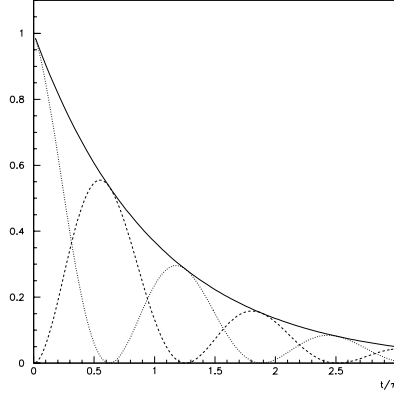
Figure 1.8: Oscillation of neutral mesons: probability of observing the original meson (dotted line), its antiparticle (dashed line) or either (solid line) as a function of elapsed proper time.

In order to detect B oscillations, it is necessary to correlate the flavour of the B meson at production with its flavor at decay time. This is done via

several **tagging** techniques discussed below. If the actual B lifetime is also known, it is possible to fit experimental data to equations 1.7 (modified to take into account possible errors in the tagging process); otherwise, only the time-integrated probability

$$\chi \equiv \frac{\mathcal{P}_{mix}}{\mathcal{P}_{mix} + \mathcal{P}_{unmix}} = \frac{x^2}{2(1 + x^2)} \qquad (1.8)$$

is available. The adimensional mixing parameter x is defined as $\Delta m/\Gamma$; the mesons oscillate $x/2\pi$ times during their average lifetime.

1.4.2   Flavour tagging

B flavour tagging consists in determining whether a B meson contains a b or a $\bar{b}$ quark. In order to study B^0 oscillation, the meson must be tagged both at production and at decay.

Many techniques are based on the study of the B meson's decay products, which are correlated to its flavour. It is possible to examine specific particles in the event, or the entire jet to which the B meson belongs.

In the semileptonic B decay $b \to \ell^- \bar{\nu} q X$, the b quark is always associated to a negatively charged lepton; decay of a $\bar{b}$ quark leads to a positive lepton. This technique is called "lepton tagging". The c quark flavour is also correlated to the b flavour, as well as its decay products; in "kaon tagging", detection of the K^- from $b \to c \to s$ identifies a b, while a K^+ is associated to a $\bar{b}$.

Other techniques rely on more than a single particle. "Jet charge tagging" algorithms, for instance, take in consideration all the charged particles in a cone around the reconstructed b direction; they correlate b flavour to a weighted sum of the charge signs of the particles in the cone. The weight function w gives more importance to the tracks that are more likely to be b decay products, rather than superimposed tracks with a different origin. Actual choices for w range from particle momentum to complicated functions of track parameters.

Another tagging method consists in examining primary vertex tracks that lie close to the B hadron. Fragmentation models, and decay modes of excited B hadrons, imply a correlation between the charge sign of these particles and the B production flavour.

The dominant B production modes described in 1.1.1 contain a $b\bar{b}$ pair in the final state. Therefore, if a B meson is detected, its flavour at production can be determined by tagging the *other* B hadron in the event, and reversing the outcome. This method is called "opposite side tagging"; methods that

provide the production flavour directly, such as fragmentation correlations, are called "same side tagging".

Tagging efficiency and dilution

Tagging is often a matter of probability; it is always possible to **mistag** a B meson, assigning to it the wrong flavour. Moreover, some tagging algorithms can only be applied in certain conditions, such as when a suitable lepton has been found. It is therefore necessary to characterize the quality of tagging.

Let us assume that, in a given sample, there are N_m^{true} events that have undergone mixing, and N_u^{true} events that have not. The tagging algorithm is able to tag a fraction ϵ of the events (tag efficiency); the probability of the tag being correct is p_R. The tagger, therefore, reports

$$
\begin{aligned}
N_m &= \epsilon \left(p_R N_m^{true} + (1 - p_R) N_u^{true} \right) \\
N_u &= \epsilon \left(p_R N_u^{true} + (1 - p_R) N_m^{true} \right)
\end{aligned}
\tag{1.9}
$$

events of each kind. The efficiency is then given by

$$
\epsilon = \frac{N_m + N_u}{N_m^{true} + N_u^{true}}
\tag{1.10}
$$

while the **dilution** is defined as

$$
\mathcal{D} \equiv p_R - (1 - p_R) = 2p_R - 1 = \frac{N_R - N_W}{N_R + N_W}
\tag{1.11}
$$

where N_R and N_W are the numbers of correct and incorrect tags. Dilution is equal to one in a perfect tagger; it is zero when the tagger is completely useless, choosing tags randomly.

In order to fit the oscillation parameters, it is necessary to estimate quantities such as the asymmetry

$$
\mathcal{A}_{true}(t) \equiv \frac{N_u^{true} - N_m^{true}}{N_u^{true} + N_m^{true}} = \cos \Delta m t
\tag{1.12}
$$

by making use of the tagger output:

$$
\begin{aligned}
\mathcal{A} &\equiv \frac{N_u - N_m}{N_u + N_m} \\
&= \frac{p_R(N_u^{true} - N_m^{true}) + (1 - p_R)(N_m^{true} - N_u^{true})}{p_R(N_u^{true} + N_m^{true}) + (1 - p_R)(N_m^{true} + N_u^{true})} \\
&= \mathcal{D} \, \mathcal{A}_{true}
\end{aligned}
\tag{1.13}
$$

where the efficiency ϵ affects numerator and denominator equally. The statistical error on $\mathcal{A}_{true}$ is

$$\sigma(\mathcal{A}_{true}) = \sqrt{\frac{1 - \mathcal{D}^2 \mathcal{A}_{true}^2}{\epsilon \mathcal{D}^2 N}} \qquad (1.14)$$

where $N \equiv N_u^{true} + N_m^{true}$ is the total sample size. When the asymmetry is small, the statistical error on N events is the same as the statistical error on $\epsilon \mathcal{D}^2 N$ perfectly tagged events; $\epsilon \mathcal{D}^2$ is therefore the figure of merit of tagging algorithms.

1.4.3   Measurements

In order to measure B oscillations, it is necessary to collect a sample of neutral B decays, to determine the proper B lifetime t, and to tag the meson as a B or $\bar{B}$. Experimental data is then fitted to the asymmetry $\mathcal{A}$ (eq. 1.13), and the mixing parameter x (or the mass difference $\Delta m = x\Gamma$) is extracted.

Once a value of x has been obtained via a maximum likelihood fit, the significance of the measurement can be defined as a number of equivalent standard deviations by the formula

$$\text{Sig} \equiv \sqrt{2\Delta \log \mathcal{L}} \qquad (1.15)$$

where $\Delta \log \mathcal{L}$ is the log-likelihood difference between the best and second-best maximum in the $\log \mathcal{L}(x)$ plot. The significance depends strongly on proper lifetime resolution σ_t:

$$\text{Sig}(x) = \sqrt{\frac{N \epsilon \mathcal{D}^2}{2}} \, e^{-\left(\frac{x^2 \sigma_t^2}{2\tau^2}\right)} \sqrt{\frac{S}{1+S}} \qquad (1.16)$$

where N is the total sample size, $\epsilon \mathcal{D}^2$ is the tagging power, S the signal to background ratio, and τ the average B meson lifetime.

The result worsens exponentially with the lifetime resolution; it is therefore essential to work on a sample where t, and hence the B momentum, can be measured accurately. As a consequence, the ideal sample for the study of B oscillations consists of events with small track multiplicity, where all the decay products are charged hadrons. Without an impact parameter trigger, however, adequate statistics and S/B ratio can only be obtained by requiring a lepton in the final state. In semileptonic decays, the neutrino momentum escapes detection, and must be inferred via Monte Carlo models. Nevertheless, good results ([27] — [31]) have been obtained for B_d oscillations at CDF using these methods; the combined CDF result is

$$\Delta m_d = (0.495 \pm 0.026(\text{stat}) \pm 0.025(\text{sys}))\text{ps}^{-1} \qquad (1.17)$$

which has a precision comparable to LEP and SLD measurements [32]:

$$
\begin{array}{ll}
\text{ALEPH} & (0.446 \pm 0.020 \pm 0.018)~\text{ps}^{-1} \\
\text{DELPHI} & (0.496 \pm 0.026 \pm 0.023)~\text{ps}^{-1} \\
\text{L3} & (0.444 \pm 0.028 \pm 0.028)~\text{ps}^{-1} \\
\text{OPAL} & (0.479 \pm 0.018 \pm 0.015)~\text{ps}^{-1} \\
\text{SLD} & (0.526 \pm 0.043 \pm 0.031)~\text{ps}^{-1}
\end{array}
$$

As for B_s mesons, the mass difference Δm_s and the mixing parameter x_s are known to be at least 30 times larger than in the B_d case. According to eq. 1.16, therefore, an actual measurement of x_s requires a σ_t improvement of the same order of magnitude. This can be done by selecting a clean sample of fully reconstructed, hadronic B_s decays; an impact parameter trigger is hence essential for the study of B_s oscillations.

1.5 CP violation

It has been known since 1957 that parity and charge conjugation symmetries are maximally broken by weak interactions; only their combination seemed to be a simmetry of the universe. In 1964, however, a slight CP breaking was observed [33] in the decay of neutral K mesons: the K_L meson, formerly thought to be the CP-odd eigenstate $\left|K_0\right\rangle + \left|\bar{K}_0\right\rangle$, was found to decay, about once every 500 times, into a CP-even two-pion state. This small asymmetry between matter and antimatter has strong cosmological consequences: it could be the reason why antimatter is so rare in the observed universe.

1.5.1 CKM matrix and unitarity triangle

Discovery of the b quark, and the consequent addition of a third fermionic generation to the Standard Model, provided evidence in favour of a theoretical argument [34] that had been proposed four years earlier to explain CP violation.

In the Standard Model lagrangian, fermion masses are introduced via the Higgs mechanism [35]. After spontaneous symmetry breaking has taken place, and once the lagrangian has been expanded around the vacuum expectation value of the Higgs field, the most general quark mass terms assume the form

$$
\bar{u}_L^a \, M_{ab}^u \, u_R^b \; + \; \bar{d}_L^a \, M_{ab}^d \, d_R^b \; + \; h.c. \tag{1.18}
$$

where a and b are generation indices, u_L and d_L form a SU(2) doublet, while u_R and d_R are singlets. The mass matrices can be diagonalized via a biunitary

transformation:

$$M_0^u = (V_L^u)^\dagger M^u V_R^u \tag{1.19}$$
$$M_0^d = (V_L^d)^\dagger M^d V_R^d$$

where the V are unitary matrices. Both V_R can be chosen to be identity matrices without loss of generality; mass and flavour eigenstates of quarks are then related to each other by the Cabibbo-Kobayashi-Maskawa matrix:

$$V \equiv (V_L^u)^\dagger V_L^d = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{1.20}$$

By redefining the quark phases, some degrees of freedom of V can be reabsorbed; with n quark generations, V contains $n(n-1)/2$ independent angles and $(n-1)(n-2)/2$ phases. For $n \geq 3$, V can contain a non-eliminable phase, which gives rise to CP violation.

The CKM matrix has been parametrized by Wolfenstein [36] in order to put in evidence the different orders of magnitude of its elements:

$$V = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4) \tag{1.21}$$

where $\lambda = 0.2196 \pm 0.0023$ is the sine of the Cabibbo angle. Since V is unitary, the relation

$$V_{ux}V_{uy}^* + V_{cx}V_{cy}^* + V_{tx}V_{ty}^* = \delta_{xy} \tag{1.22}$$

holds true for any pair of columns x, y. If $x \neq y$, the three terms form a closed triangle in the complex plane — the CKM unitarity triangle of figure 1.9. All the possible triangles have the same area, which is proportional to the intensity of CP violation; the triangle made with the d and b columns is the least flattened, which makes it easier to measure its angles. If the triangle is appropriately rescaled, and two of its vertices are placed in $(0,0)$ and $(1,0)$, the third vertex's coordinates are the Wolfenstein parameters ρ and η.

The unitarity triangle can be constrained with a wide range of analyses, including independent measurement of its sides and angles. Should the triangle be found not to close exactly, there would be indirect evidence for phenomena beyond the Standard Model, such as a fourth generation of fermions.
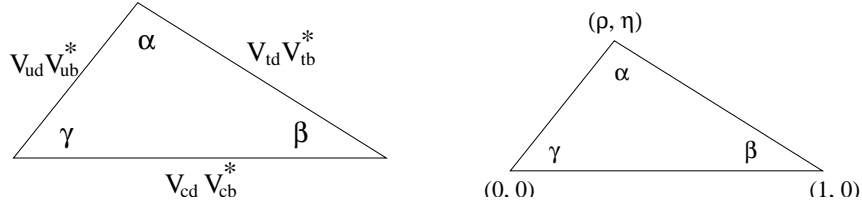
Figure 1.9: Unitarity triangle as a function of V_{CKM} matrix elements and in the Wolfenstein parametrization

At present, the best results constrain the triangle sides and force the vertex to lie within a certain band of the complex plane. Improvement on these measurements is limited by QCD uncertainties. On the other hand, direct measurements of the angles are not yet sensitive enough to improve the situation; this will change in the near future.

1.5.2 Measurement of CKM angles

B^0 mesons can decay into the same CP eigenstate $J/\psi\, K_S^0$ in two different ways: either directly, or after oscillating into $\bar{B}^0$. The interference between these diagrams causes a decay asymmetry between the mesons that originally contained a b or a $\bar{b}$ meson:

$$\mathcal{A}_{CP}(t) \equiv \frac{\bar{B}^0(t) - B^0(t)}{\bar{B}^0(t) + B^0(t)} = \sin 2\beta \cdot \sin \Delta m_d t \qquad (1.23)$$

where $B^0(t)$ and $\bar{B}^0(t)$ are the numbers of mesons that were B (or $\bar{B}$) at production and decayed at proper time t. This asymmetry oscillates with the same period as $B \leftrightarrow \bar{B}$ oscillations, but with a $\pi/2$ phase shift.

This measurement has been attempted at the CDF experiment [37]. The analysis combined several tagging techniques, applying them to data samples of different quality (with or without precise vertex measurement). The tagging power $\epsilon\mathcal{D}^2$ of the four tagging methods ranged from $(1.0 \pm 0.3)\%$ to $(2.2 \pm 1.3)\%$; besides, the sample size was small (400 events overall, 200 with clear vertex information). The analysis was based on an unbinned likelihood fit, in order to make optimal use of the statistics; its result was

$$\sin 2\beta = 0.79 \pm 0.36(\text{stat}) \pm 0.16(\text{sys}) \qquad (1.24)$$

where the large statistical error is due to the small sample and tagging power, while the systematic error is dominated by the uncertainty on $\mathcal{D}$. For a more accurate $\sin 2\beta$ measurement at CDF, it is therefore essential to increase the

sample size and to provide a better estimate of the dilution. As a comparison, a recently presented preliminary result from BaBar [38] states

$$\sin 2\beta = 0.12 \pm 0.37(\text{stat}) \pm 0.09(\text{sys}) \qquad (1.25)$$

Another CKM measurement, also based on the observation of B decays into a CP eigenstate, involves a study of the interference between $B^0 \to \pi^+\pi^-$ and $B^0 \to \bar{B}^0 \to \pi^+\pi^-$. As in the previous case, the asymmetry [39]

$$\mathcal{A}_{CP}(t) \equiv \frac{\bar{B}^0(t) - B^0(t)}{\bar{B}^0(t) + B^0(t)} \simeq \sin 2\alpha \cdot \sin \Delta m_d t \qquad (1.26)$$

oscillates as a function of proper time t; its amplitude is related to $\sin 2\alpha$.

Collecting a $B \to \pi\pi$ sample is not a simple task, especially at a hadronic collider: the final state does not contain any leptons on which to trigger, and the events of interest have to be extracted from a very large background. Moreover, the $B \to \pi\pi$ branching ratio is quite small, of the order of $4.3 \, 10^{-6}$ [40]. In order to achieve a good result, it is necessary to select these events at run time, triggering on the high impact parameter of the decay products. The experiment trigger must therefore be able to perform accurate track reconstruction in a very short time. Particle identification, via invariant mass reconstruction, time of flight and rate of energy loss, is also important: the signal must be separated from $B \to \pi K$ and $B \to KK$ backgrounds.

1.6   CDF II: expectations and requirements

At present, there are several open questions in the field of B physics. Some elements of the CKM matrix are still little known, and the unitarity triangle is not yet constrained; the x_s mixing parameter for B_s flavour oscillations is beyond the reach of currently available data samples.

In the past years, Fermilab and the CDF collaboration have planned and implemented an accelerator and detector upgrade, in order to give better answers to these and other questions. CDF II is now undergoing the final phases of assembly and testing; it is scheduled to become operative in March 2001.

From an extrapolation of previous results to the new environment, CDF II is expected — for instance — to be able to measure $\sin 2\beta$ with an accuracy of ± 0.08, and $\sin 2\alpha$ to within ± 0.10. As shown in figure 1.10, it will be sensitive to B_s oscillations over a wide range of x_s values, starting from the lower bound set by current experimental results, and extending up to values of 50–60. CDF II will also be able to continue the study of B_c mesons, which
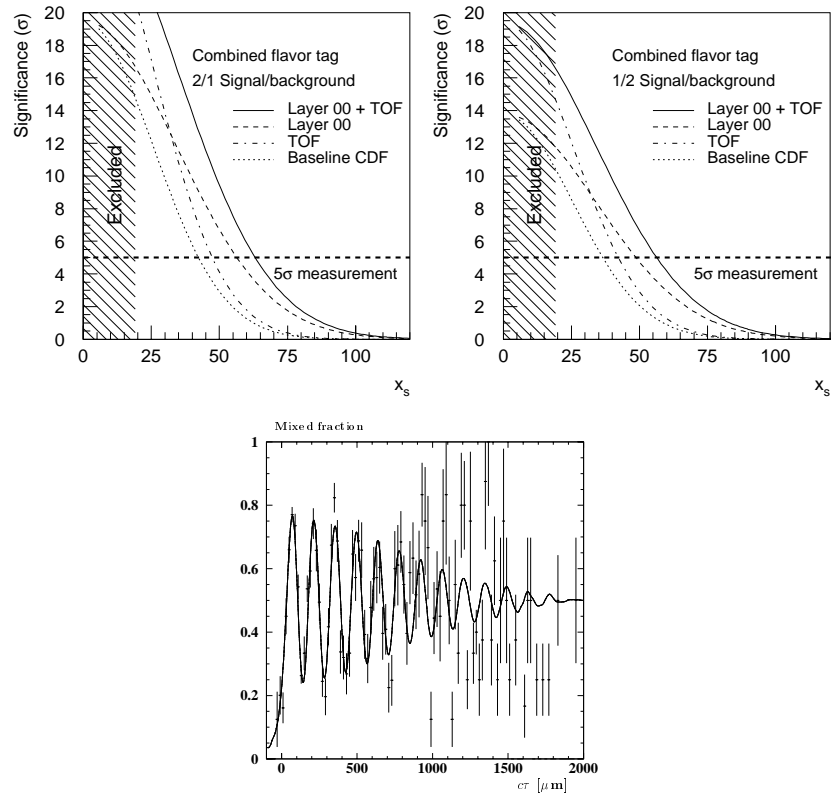
Figure 1.10: B_s oscillations at CDF II: x_s significance as a function of x_s, in the case of 2:1 (left) and 1:2 (right) signal to noise ratio; simulation results (below) for $x_s = 20$.

are also not produced at $\Upsilon(4S)$ B factories, and the search for rare B decay modes.

In order to achieve these results, the experiment must satisfy several requirements. First of all, it must be able to collect large samples of rare B decays, in order to decrease statistical errors. The accelerator must therefore produce a sufficient number of collisions; the detector must reconstruct as many signal events as possible, while rejecting the background.

To increase the signal sample, the detector's geometric acceptance must be as wide as possible. Measuring the proper lifetime of a B meson requires a good knowledge of the particle's momentum; this can be reconstructed precisely only if all the decay products have been detected. Therefore, even a modest increase in the acceptance region can have large repercussions on the overall efficiency.

Another essential element of b analyses is the measurement of vertex positions. The experiment must detect the position of tracks as close as possible to the interaction point, with an excellent resolution; at the same time, the amount of material in the detector must not degrade track quality via excessive multiple scattering.

Particle identification also plays a fundamental role in B physics. Analyses of semileptonic B decays rely on the charged lepton as a signature; in other cases, such as in $\sin 2\alpha$ measurements, it is necessary to distinguish between π and K, by examining the particle's dE/dx and time of flight.

Finally, the trigger system must be able to provide high efficiency and high rejection rate; at the same time, it must be fast enough to take advantage of the high collision frequency of the accelerator, without introducing unnecessary dead time.

The CDF II experiment satisfies all of these requirements: the following two chapters will describe in detail the technical choices that were made to achieve the experiment's physics goals.

Chapter 2

Tevatron and CDF Upgrades

2.1 A short history of CDF

The CDF experiment (Collider Detector at Fermilab) [41] aims at studying collisions between protons and antiprotons at the Tevatron accelerator, at center-of-mass energies up to $\sqrt{s} = 2\,\mathrm{TeV}$.

The first events were detected in 1985; during the subsequent eleven years, the increasing luminosity of the accelerator and several improvements in the detector led to the accumulation of larger and more sensitive data samples:

	1987	25 nb^{-1}
Run 0	1988-1989	4.5 pb^{-1}
Run 1a	1992-1993	19 pb^{-1}
Run 1b	1994-1996	90 pb^{-1}

Analyses of experimental data resulted in the publication of more than 170 papers [42], ranging over the entire spectrum of hadron collider physics. To cite but a few results, CDF provided the first experimental evidence for the top quark [43][44][45] and accurate measurements [46] of its mass, $m_t = 176.1 \pm 6.6\,\mathrm{GeV}/c^2$; precision electroweak measurements, such as [47] $m_W = 80.433 \pm 0.079\,\mathrm{GeV}/c^2$; and determination of average lifetime for several b hadrons [2].

Since the shutdown in 1996, the Tevatron and its detectors — CDF and D0 — have undergone major upgrades, in order to make them able to provide the results discussed in Chapter 1 during the physics run that is scheduled to begin in March 2001. The rest of this chapter describes the various upgrades, and their impact on the physics program.

2.2   The Accelerator Complex

In the next decade, CDF plans to carry out precise analyses of several rare physical processes whose cross section is several orders of magnitude smaller than the inelastic $p\bar{p}$ cross section. In order to obtain sufficiently large samples, several steps have been taken:

- Increase the center-of-mass energy

- Increase the luminosity

- Increase the detector's acceptance

The first two steps, and the partial reconstruction of the Tevatron which they implied, are the topic of this section.

As was stated above, the Run II proton-antiproton center of mass energy will increase to 2 TeV from the Run I value of 1.8 TeV. While this change will affect the performance of the detectors only marginally[1], it will provide a major increase in the reconstructed sample size; for example, the cross section for associated $t\bar{t}$ production will grow by 40% with respect to Run I.

Another way to obtain a larger sample is to increase the accelerator's luminosity. In the ideal case, where the proton and antiproton beams collide head-on without a crossing angle and with optimal alignment, the Tevatron's luminosity [48] is given by the formula

$$\mathcal{L} = \frac{f B N_p N_{\bar{p}}}{2\pi(\sigma_p^2 + \sigma_{\bar{p}}^2)} F\left(\frac{\sigma_l}{\beta^*}\right) \qquad (2.1)$$

where f is the revolution frequency, B the number of bunches in each beam, N_p and $N_{\bar{p}}$ the number of protons and antiprotons per bunch, σ_p and $\sigma_{\bar{p}}$ the transverse beam sizes (RMS) at the interaction point, and F a form factor that depends on the ratio between the bunch longitudinal RMS size, σ_l, and the beta function[2] at the interaction point, β^*.

As shown in Table 2.1, the most significant improvements in luminosity will be obtained by increasing the number of bunches per beam from 6 to 36, and then to 108, while keeping the number of particles per bunch similar to or higher than the Run I figure.

---

[1]As will be shown in Chapter 3, the detector's performance depends mainly on the number of superimposed events, which will be discussed shortly.

[2]Supposing the profile of the beam in the phase space (x, x') is an ellipse of semiaxes σ and σ', the *amplitude function* β is defined as the ratio σ/σ', while the beam *emittance* is the phase volume $\epsilon = \pi\sigma\sigma'$.

Run	1989	IA (1992-93)	IB (1993-95)
p/bunch	7.00E+10	1.20E+11	2.32E+11
$\bar{p}$/bunch	2.90E+10	3.10E+10	5.50E+10
p emittance (mm mrad)	25	20	23
$\bar{p}$ emittance (mm mrad)	18	12	13
Beta @ IP (m)	0.55	0.35	0.35
Energy (GeV/particle)	900	900	900
Bunches	6	6	6
Bunch length (rms, m)	0.65	0.55	0.6
Form Factor	0.71	0.62	0.59
Typical $\mathcal{L}$ (cm^{-2}s^{-1})	1.60E+30	5.42E+30	1.58E+31
Best $\mathcal{L}$ (cm^{-2}s^{-1})	2.05E+30	9.22E+30	2.50E+31
$\int \mathcal{L}dt$ (pb^{-1}/week)	0.32	1.09	3.18
Bunch Spacing (nsec)	3500	3500	3500
Interactions/crossing	0.25	0.85	2.48
What's New?		Separators $\bar{p}$ improvements	Linac Upgrade

Run	II (2001-)	II (2001-)	TEV33
p/bunch	3.30E+11	2.70E+11	2.40E+11
$\bar{p}$/bunch	3.60E+10	5.50E+10	1.00E+11
p emittance (mm mrad)	30	18	18
$\bar{p}$ emittance (mm mrad)	20	15	18
Beta @ IP (m)	0.35	0.35	0.25
Energy (GeV/particle)	1000	1000	1000
Bunches	36	36	108
Bunch length (rms, m)	0.43	0.18	0.26
Form factor	0.70	0.90	0.75
Typical $\mathcal{L}$ (cm^{-2}s^{-1})	8.29E+31	2.03E+32	1.04E+33
Best $\mathcal{L}$ (cm^{-2}s^{-1})	N/A	N/A	N/A
$\int \mathcal{L}dt$ (pb^{-1}/week)	16.72	40.83	210.62
Bunch Spacing (nsec)	396	396	132
Interactions/crossing	2.17	5.31	9.13
What's New?	Main Injector $\bar{p}$ improvements	New $\bar{p}$ ring	Targeting Cooling upgr.

Table 2.1: Evolution of Tevatron parameters. "Typical" luminosity is quoted at the beginning of a store; it translates to integrated luminosity with a 33% duty factor.

A limiting factor in the choice of accelerator parameters is the superposition of multiple elementary proton-antiproton interactions within the same bunch crossing. At high luminosities, this superposition increases the complexity of the event, making its reconstruction more difficult.

The actual number of superimposed interactions obeys a Poisson distribution, whose mean (as a function of luminosity) is shown in Figure 2.1 for different values of the number of bunches.
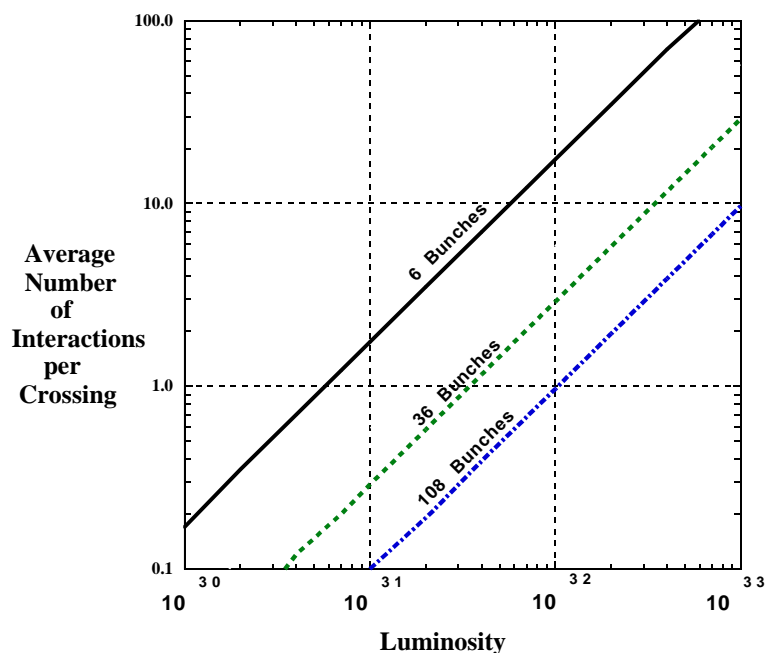

Figure 2.1: Average number of $p\bar{p}$ interactions per bunch crossing, as a function of luminosity and of the number of bunches.

Production and acceleration of protons and antiprotons at Fermilab requires a chain of accelerators, each boosting particles to higher energies. Each step will be described in the following pages.

2.2.1 Proton production and boosting

The process begins with a **Cockcroft-Walton** accelerator, which feeds negative hydrogen ions to a 150 m linear accelerator. The Linac itself was upgraded in 1993, increasing its energy from 200 MeV to 400 MeV; this made it possible, during Run Ib, to double the number of protons per bunch, and to increase by about 50% the production rate of antiprotons.

After being stripped of electrons, the protons enter the **Booster**, a synchrotron whose diameter is about 150 m, where they reach a kinetic energy of 8 GeV. Together, Linac and Booster will be able to provide pulses of $5 \cdot 10^{12}$ protons for antiproton production every 1.5 s, or $6 \cdot 10^{10}$ protons per bunch in series of 5 to 7 bunches, repeated 36 times every four seconds.

After leaving the Booster, protons are transferred to the Main Injector, a newly built circular accelerator that will replace the older Main Ring.

2.2.2 Main Injector

The Main Ring was originally built to provide 400 GeV protons to Fermilab's fixed target experiments; later on, it was converted to act as an injector to the Tevatron. The new operational requirements for the Main Ring did not match its original design; therefore, during Run I, the Main Ring was a performance bottleneck. To quote an example, the Main Ring was never able to make full use of the Booster's capabilities: the Main Ring's aperture $(12\pi$ mm mrad$)^3$ is only 60% of the Booster's aperture $(20\pi$ mm mrad). The situation would be even worse in Run II, with the Booster's aperture at injection increasing to 30π mm-mrad.

The **Main Injector** was designed to solve this problem, while providing further benefits. It is a 3-km circular accelerator, which brings protons and antiprotons from a kinetic energy of 8 GeV to a total energy of up to 150 GeV. Its transverse admittance is larger than 40π mm mrad, more than enough to accommodate particle bunches from the Booster; its emittance is about 12π mm mrad. The maximum beam size is $3 \cdot 10^{13}$ particles, divided into up to 504 bunches of $6 \cdot 10^{10}$ (anti)protons.

Being more flexible than the Main Ring, the Main Injector can be used in several operation modes:

- Antiproton production;

- Proton and antiproton boosting, before injection into the Tevatron in collider mode;

- Antiproton deceleration, in order to recover unused antiprotons after a Tevatron collision run;

- Proton and antiproton acceleration for fixed target experiments, either directly or as a booster for the Tevatron.

---

[3]All emittances are normalized at 95% of the beam.

2.2.3   Antiproton production

In order to produce antiprotons, a pulse of $5 \cdot 10^{12}$ protons at 120 GeV is extracted from the Main Injector and focused on a nickel target. A lithium lens collects the antiprotons produced by the collision, with a wide acceptance around the forward direction, at energies close to 8 GeV. The antiproton bunches are then moved to a **Debuncher Ring**, where they are transformed into a continuous beam and stochastically cooled, and then to the **Accumulator**, where they are further cooled. The antiproton stacking rate during Run I was about $7 \cdot 10^{10}$ $\bar{p}$/hour; Run II upgrades, ranging from antiproton cooling to improving the lithium lens, will increase this rate by a factor of three to four.

When a sufficient number of antiprotons (up to 10^{12}) is available, stacking is suspended; the antiprotons are further cooled, and then transferred, with an aperture of 10π mm mrad and a $\Delta p/p < 10^{-3}$, to the antiproton Recycler Ring.

2.2.4   Recycler Ring

The **Recycler Ring** lies in the same enclosure as the Main Injector; contrarily to the other rings at Fermilab, it is built with permanent magnets. During Run I, the antiproton accumulation ring was found to suffer some kind of failure approximately once a week; this led to the loss of the entire store. Permanent magnets, not being prone to the most common causes of failure (such as power loss and lightning) will provide a very stable repository for up to $3 \cdot 10^{12}$ antiprotons at a time.

During Run II, bunches of $2 \cdot 10^{11}$ recently produced antiprotons will be transferred from the Accumulator to the Recycler Ring every about half an hour, thus keeping the total beam current in the Accumulator small (below 10 mA, compared to the 200 mA antiproton current in Run I).

After a Run I collider store was over, the antiprotons remaining in the collider used to be dumped. In Run II, instead, they will be slowed down by the Tevatron and Main Injector, back to an energy of 8 GeV; they will then be stored in the Recycler Ring, subject to electron cooling, and reused in the following run. This feature earns the Recycler Ring its name.

Antiproton production is one of the limiting factors in the efficiency of Fermilab's colliders. At the end of a store, 75% of the antiprotons are expected to be still circulating in the Tevatron; by recycling 2/3 of these antiprotons, the average luminosity can be increased by a factor of two.

2.2.5 Tevatron

The Tevatron is a 6-km circular accelerator, where protons and antiprotons, rotating in opposite directions inside the same beam pipe, are accelerated from 150 GeV to 1 TeV. Making use of the upgrades in the rest of the accelerator chain, the Tevatron will be able to provide an initial luminosity of $5 \cdot 10^{31} \mathrm{cm}^2 \mathrm{s}^{-1}$, which will increase up to $2 \cdot 10^{32} \mathrm{cm}^2 \mathrm{s}^{-1}$ after fine-tuning. In this operation mode, the Tevatron will be able to provide 2 fb^{-1} of data within two years.

During a collider store, instant luminosity slowly decreases. In the early stages of the store, the most important cause for this decrease is intrabeam scattering; some hours later, the depletion of antiprotons during collisions becomes more relevant. Luminosity is expected to decrease to 50% in about seven hours, and to $1/e$ in twelve hours. After a typical store duration of eight hours, 75% of the antiprotons are still available; they are decelerated in the Tevatron and in the Main Injector, and then stored in the Recycler Ring and re-cooled.

The Tevatron can also be used in fixed-target mode: it can accelerate up to $3 \cdot 10^{13}$ protons at a time to an energy of 800 GeV, and deliver single bunches to be used in proton, meson and neutrino experiments.

Other operational parameters of the Tevatron are listed in Table 2.1.

2.2.6 Beam monitors

Operation of colliders at the Tevatron requires a constant monitoring of the beam position and luminosity. From a conceptual point of view, this is done in Run II as it was done in Run I.

The luminosity monitor consists in two arrays of scintillators, placed on both sides of the interaction region. A coincidence of particles moving away from the interaction point, both in the p and $\bar{p}$ direction, is interpreted as a contribution to luminosity; bunches of particles moving in a single direction, without a coincident bunch in the opposite direction, are flagged as beam losses.

The beam position, on the other hand, is measured by the collider detectors themselves. During Run I, the detector was able to locate the beam within 5 μm in about five minutes; other beam parameters, such as slope and transverse profile, were calculated over longer time intervals (about two hours). In Run II, the same operations will be performed more quickly, thanks to the faster electronics and the higher collision rate.

2.3   CDF II: Overview

As stated above, one of the aims of Run II is to reconstruct and store a large
sample of rare events. To achieve this result, the number of bunches in each
beam will increase first by a factor of six with respect to Run I, and then by a
further factor of three. An immediate consequence is that the time between
two successive interactions will decrease by the same factor, from $2.4\,\mu s$ to
400 ns and then to 132 ns. Several parts of the detectors have been rebuilt
from scratch in order to accomodate the higher collision rate.

While the detector was redesigned [49], efforts were also made to extend
its acceptance. The geometrical coverage was increased, by adding new de-
tector elements or enlarging the previously existing ones; the trigger system
became able to detect some interesting event features at an earlier stage than
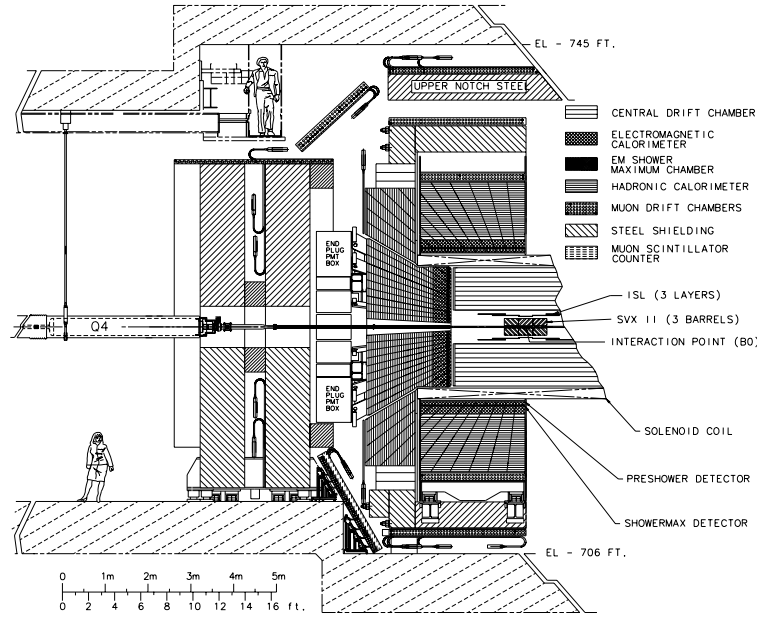in Run I, thus improving the signal to background ratio.

Figure 2.2: Elevation view of one half of the CDF II detector

As shown in figure 2.2, the tracking system of CDF II is placed inside a
superconducting solenoid, while calorimeter and muon systems are outside
the magnet. The rest of this chapter will provide a short description of the
detector subsystems, with an emphasis on the upgrades since Run I; the
tracker will be examined in deeper detail in Chapter 3.

In the standard CDF geometry, the $\hat{z}$ axis is oriented along the axis of

the solenoid, the $\hat{x}$ axis points away from the center of the Tevatron, and the $\hat{y}$ axis points up. The origin is at the interaction point. The polar angle θ is measured starting from the positive $\hat{z}$ axis; the pseudorapidity is defined as

$$\eta = -\log\left(\tan\frac{\theta}{2}\right) \qquad (2.2)$$

2.4 Tracking system

The innermost parts of the CDF II detector are devoted to tracking charged particles. The tracking system will be examined in depth in Chapter 3; a shorter description is given here.

2.4.1 Silicon Vertex Detector

CDF II makes use of three concentric silicon detectors: "Layer 00" (**L00**), the Silicon Vertex Detector (**SVX II**, or **SVX** in short), and the Intermediate Silicon Layers (**ISL**).

SVX II is the Run II baseline detector. It consists of five layers of double-sided silicon wafers. One side of each wafer provides measurements in the transverse plane (axial strips); the other side's strips deliver 3D information. SVX II extends radially from 2.5 to 10 cm, and along z up to 45 cm on either side of the interaction point.

The ISL consists of a double-sided silicon layer, similar to those in SVX II, placed at $r = 22$ cm in the central η region, and of two forward layers ($1 < |\eta| < 2$) respectively at 20 and 29 cm from the beam line. Together with SVX II, the ISL makes it possible to reconstruct tracks in the forward region, which lies beyond the acceptance region of the outer tracker.

Layer 00 is the most recent addition to the CDF II tracker. It is a single sided, radiation–hard silicon layer, placed immediately outside the beam pipe, at $r \simeq 1.5$ cm. Being so close to the interaction point, Layer 00 improves noticeably the impact parameter resolution. In case the innermost SVX II layer suffers from radiation damage during Run II, Layer 00 will also act as a backup.

Compared to the shorter, 4-layer, single-sided vertex detector of Run I, the new silicon tracker provides a much wider acceptance, better resolution, three-dimensional reconstruction, and can be used in stand-alone mode, without input from the Central Outer Tracker (described hereafter).

2.4.2   Central Outer Tracker

Outside the silicon detector, at a distance between 40 and 138 cm from the beam, lies the Central Outer Tracker. It is a new open-cell drift chamber, able to reconstruct tracks in the $|\eta| < 1$ region, provided their momentum is at least 300 GeV/c. The COT replaces an older drift chamber, the CTC [50], that would have been unable to cope with the expected occupancy and event rate of Run II.

Each of the eight superlayers of cells consists of twelve layers of sense wires, alternating with field-shaping wires. Axial superlayers alternate with stereo superlayers, thus providing 48 axial and 48 stereo measurements for each track.

In the COT, the cell size is roughly four times smaller than in the CTC. Usage of a faster gas ($Ar - Ethane - CF_4$ instead of $Ar - Ethane$) reduces the maximum drift time by a further factor of two, down to 100 ns. This makes the COT immune from event pile-up, even at the highest collision rate of $1/(132$ ns$)$.

2.4.3   Time of flight

A recent addition to CDF II, the time-of-flight detector is an array of scintillator bars, placed at the outer edge of the COT, at a radial coordinate of 140 cm. As plotted in figure 2.3, an accurate measurement of a particle's time of flight in the CDF tracking volume can be used quite effectively in particle identification; this alone can improve B_s tagging power by a factor of two, with the consequences on x_s sensitivity shown in figure 1.10.

Scintillator bars are about three meters long, matching the COT active volume; their thickness (4 cm) is limited by the space which remained available between the previously designed COT and magnet. Their width was determined by occupancy[4] and resolution considerations; the best choice turned out to be also of the order of 4 cm. The bars have a trapezoidal cross section, in order to minimize cracks in the geometry; the scintillating material is Bicron 408, which has a short rise time and a long (380 cm) attenuation length.

Photomultiplier tubes, attached to both ends of each bar, provide time and pulse height measurements. By comparing the two pairs of results, the detector determines the instant in which a particle crossed the scintillator with an accuracy of about 100 ps, and the z coordinate of the intersection.

---

[4]Detector occupancy depends on the average number of superimposed interactions, which increases with luminosity (fig. 2.1). TOF occupancy is estimated to be 0.1 with 2 superimposed events, and 0.4 with 10 events.
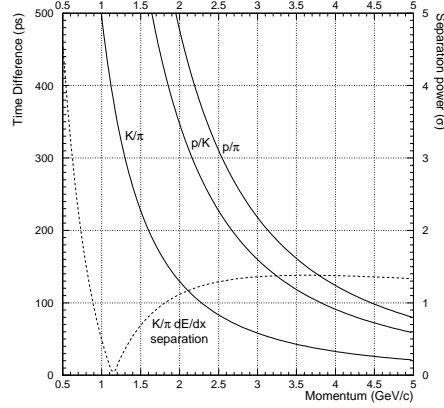
Figure 2.3: Time of flight difference as a function of particle type and momentum. The dashed line indicates the COT dE/dX separation power for $K\pi$: TOF and COT complement each other in different p regions.

The latter measurement is compared to the results of 3D track reconstruction in the inner tracking volume, to associate a time of flight to each track.

2.4.4   Magnet

The CDF tracking systems are enclosed in a superconducting solenoid, which provides a uniform magnetic field of up to 1.5 T along the detector axis, over a cylindrical fiducial volume 3.5 m long and 2.8 m in diameter.

The solenoid is built of an Al-stabilized NbTi superconductor, able to withstand currents up to 5000 A, and operating at liquid helium temperature. During most of Run I, the magnet operated at 4650 A, corresponding to a current density of 1115 A/m and a central field of 1.41 T.

Although the design lifetime of the solenoid was only ten years, it will be possible to reuse the magnet during Run II. The cool-down procedures that were used during Run I limited mechanical stress to the coil, avoiding fatigue damage.

2.5   Calorimetry

2.5.1   Overview

CDF uses scintillator sampling calorimeters, divided into separate electromagnetic and hadronic sections, and providing coverage for $|\eta| \leq 3.64$. The

calorimeter was an essential tool in selection and reconstruction of events in
Run I; in Run II it will continue to measure the energy of photons, electrons,
jets, and the missing transverse energy [5] associated to neutrinos and possibly
to neutral exotic particles.

Calorimeter calibration can be performed by matching the tracks found in
the tracking system to the corresponding calorimetry towers; during Run I,
this provided a 2.5% accuracy on jet energy measurements.

$|\eta|$ range	$\Delta\phi$	$\Delta\eta$
0 — 1.1 (1.2 had)	15°	0.1
1.1 (1.2 had) — 1.8	7.5°	0.1
1.8 — 2.1	7.5°	0.16
2.1 — 3.64	15°	0.2 — 0.6

Table 2.2: Calorimeter segmentation

The entire calorimeter is segmented into projective towers, whose geom-
etry is summarized in table 2.2. Each tower consists of alternating layers
of passive material (lead for the e.m. section, iron for the hadronic com-
partment) and scintillator tiles. The signal is read via wavelength shifters
(WLS) embedded in the scintillator; light from the WLS is then carried to
photo-multiplier tubes. Table 2.3 shows the most important characteristics
of each calorimeter sector. The central and end-wall calorimeters ($|\eta| < 1.1$)
[51] [52] were recycled from Run I; the plug ones ($1.1 < |\eta| < 3.64$) were built
anew, to replace an older gas calorimeter that would not be able to function
at the increased event rate of Run II.

2.5.2   Central calorimeter

Apart from the electronics, the central calorimeter in CDF Run II will be
the same used during Run I. The energy measurement response time is al-
ready fast enough to accomodate a 132 ns bunch spacing; the preshower wire
chambers will require integration over a few consecutive events, but their
granularity is fine enough not to make this a problem.

---

[5] Contrarily to e^+e^- colliders, in $p\bar{p}$ colliders the longitudinal momentum of the initial
state is unknown.

	Central and End–wall	Plug
Electromagnetic:		
Thickness	$19\ X_0, 1\ \lambda$	$21\ X_0, 1\ \lambda$
– per sample (Pb)	$0.6\ X_0$	$0.8\ X_0$
– per sample (scint.)	5 mm	4.5 mm
Light yield	160 p.e./GeV	300 p.e./GeV
Sampling resolution	$11.6\% / \sqrt{E}$	$14\% / \sqrt{E}$
Stochastic resolution	$14\% / \sqrt{E}$	$16\% / \sqrt{E}$
Hadronic:		
Thickness	$4.5\ \lambda$	$7\ \lambda$
– per sample (Fe)	1 in (central)	2 in
	2 in (end–wall)	
– per sample (scint.)	6 mm	6 mm
Light yield	40 p.e./GeV	39 p.e./GeV
Resolution	$75\%/\sqrt{E} \oplus 3\%$	$80\%/\sqrt{E} \oplus 5\%$

Table 2.3: Characteristics of the CDF II calorimeter

Central electromagnetic calorimeter

The central electromagnetic calorimeter consists of projective towers of alternating lead and scintillator. The signal is read via a PMMA[6] wavelength shifter, and carried via clear fiber to photomultiplier tubes. None of these is expected to suffer much from radiation damage. The light yield loss is expected to be around 1% per year; 60% of this loss is explained by the gradual shortening of the attenuation length in the scintillator.

A two-dimensional wire chamber is embedded in the calorimeter, as a shower maximum detector. Its usage in the Run I trigger decreased the fake electron trigger rate by a factor of two [53]; this performance is expected to improve during Run II.

Another wire chamber is placed immediately in front of the calorimeter, to act as a pre-shower detector (CPR) which uses the tracker and the solenoid coil as radiators. The CPR has proven to be extremely useful in rejection of electron background; it also reduced systematic uncertainties for direct photon measurements by a factor of three [54].

---

[6]PMMA = polymethylmethacrylate

Central hadronic calorimeter

The central and end-wall hadronic calorimeters use 23 iron layers as radiator. The scintillator should not suffer radiation damage from measured events; however, it will require shielding from the beam halo.

The hadronic compartment geometry matches the projective towers of the electromagnetic calorimeter.

2.5.3   Plug calorimeter upgrade

The CDF II plug calorimeter, shown in figure 2.4, covers the η region between 1.1 and 3.64, corresponding to polar angles between 37° and 3°. It replaces an older gas calorimeter, whose response speed was too slow for usage at the CDF II 132 ns interbunch. Being based on the same principles as the central calorimeter, the new plug calorimeter will also make experimental data more homogeneous.
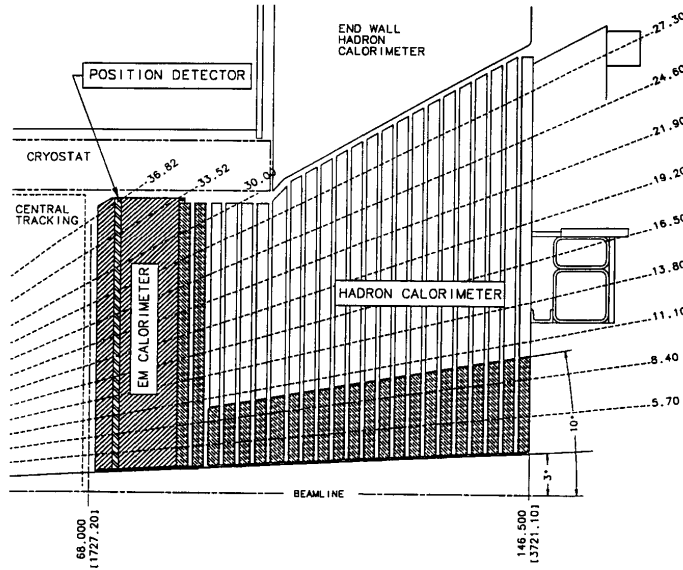

Figure 2.4: Cross section of half CDF II plug calorimeter

The calorimeter is divided in 12 concentric η regions, which are further segmented in 24 (for $|\eta| < 2.11$) or 12 (for $|\eta| > 2.11$) projective towers. The actual size of these towers was chosen so that the identification of electrons in b jets would be optimized; figure 2.5 shows the best-case identification probability as a function of η.
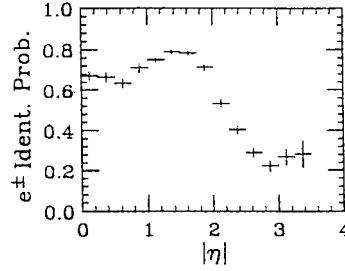
Figure 2.5: Identification probability for electrons from $b \to e + X$

Plug electromagnetic calorimeter

The EM section of the plug calorimeter consists of 23 absorber-scintillator layers. A calcium-tin-lead alloy, enclosed between steel plates, is used as absorber.

The first layer of the EM section is used as a preshower detector. In order to distinguish γ from π^0 reliably, the light yield needs to be higher than on other layers. Therefore, the first scintillator layer is thicker (10 mm instead of 6 mm) and made of a brighter material; it is read out separately from the rest of the calorimeter, via multi-anode photomultiplier tubes (MAPMT).

As in the central calorimeter, a shower maximum detector (SMD) is also embedded in the plug EM calorimeter, at a depth of about six radiation lengths. The plug SMD consists of eight 45° sectors, each covering six (or three) calorimetric towers in ϕ; each sector is further segmented in two η regions, in order to reduce occupancy. Within each region, scintillating strips are arranged on two layers, in directions parallel to either edge of the sector; this provides a two-dimensional measurement of the shower. The strips are 5 mm wide and 6 mm thick; they are read out via WLS fibers and MAPMT.

The SMD will be used to measure the position of electromagnetic showers with an accuracy reaching 1 mm for high-energy electrons, and to discriminate pions from photons and electrons.

Plug hadronic calorimeter

The hadron plug calorimeter was designed to optimize detector performance on b, electroweak and jet physics, and to help in muon detection by analyzing their rate of energy loss. It will achieve an energy resolution of about $80\%/\sqrt{E} \oplus 5\%$, which is dominated by the sampling fluctuations from the steel absorber plates. The most strict requirement is that the light yield within each tile should be uniform to 4% or better; disuniformity between

different tiles is not as important, as the hadron shower usually affects ten
or more layers.

2.6   Muon chambers

The outermost component of CDF II is a set of scintillators, drift tubes and
steel absorbers, used for the detection of muons.

During Run I, detection of muons has proven to be an important require-
ment, both for the analysis of several physics channels and for calibration.
For example, a clean sample of W bosons is obtained by reconstructing their
muon decay mode; $J/\psi \to \mu^+\mu^-$ decays are an important part of the heavy
quark physics program, as well as a tool to measure systematic effects in the
detector.

The tracking improvements from Run I to Run II have a deep impact on
muon detection. Before the upgrades, muons in the central region were iden-
tified by their penetrating power (they would not be stopped in the calorime-
ter), and their momentum was measured in the central tracking chamber. On
the contrary, the momentum of forward muons had to be measured in the
muon chambers themselves, by resorting to a toroidal magnet, as the central
tracker only covered the $|\eta| < 1$ region.

With the SVX II upgrade, this distinction falls: measurement of muon
momentum will always be performed in the central tracker, where the mul-
tiple scattering effects are smaller, and the toroidal magnets will not be
required any longer. Central tracks will be measured in the drift chamber;
forward tracks ($|\eta| > 1$) will be tracked in the silicon only.

During Run II, the Run I central muon chambers (**CMU**) will be reused
without major changes; some upgrades which started under Run I (**CMP**
and **CSP**, the Central Muon/Scintillator Upgrades; **CMX** and **CSX**, the
Central Muon/Scintillator Extension) will be completed; and a new set of
chambers, the Intermediate Muon Detector **IMU**, will replace the previous
Forward Muon Detectors (**FMU**)[55].

Due to their size, muon systems will be unable to take data within the
Run II interbunch interval of 400 or 132 ns; this is not a problem, since the
low occupancy of the muon chambers allows integration over multiple events.
Scintillators will be used to associate muon stubs to the appropriate event.

Table 2.4 summarizes the information on the muon subsystems; the fol-
lowing sections will describe their characteristics in deeper detail.

	CMU	CMP/CSP	CMX/CSX	IMU
η coverage	0 — 0.6	0 — 0.6	0.6 — 1.0	1.0 — 1.5
Drift tubes:				
thickness	2.68 cm	2.5 cm	2.5 cm	2.5 cm
width	6.35 cm	15 cm	15 cm	8.4 cm
length	226 cm	640 cm	180 cm	363 cm
max drift time	0.8 μs	1.4 μs	1.4 μs	0.8 μs
# tubes (Run Ib)	2304	864	1536	—
# tubes (Run II)	2304	1076	2208	1728
Scintillators:				
thickness	N/A	2.5 cm	1.5 cm	2.5 cm
width	N/A	30 cm	30 - 40 cm	17 cm
length	N/A	320 cm	180 cm	180 cm
# counters (Run Ib)	N/A	128	256	—
# counters (Run II)	N/A	269	324	864
π^0 int. lengths	5.5	7.8	6.2	6.2 — 20
Min P_t (GeV/c)	1.4	2.2	1.4	1.4 — 2.0
MS resol. (cm GeV)	12	15	13	13 — 25

Table 2.4: Parameters of muon detection at CDF. Pion interaction length and the limit on resolution due to multiple scattering are computed at $\theta = 90°$ in the central detectors CMU, CMP and CSP; at $\theta = 55°$ in CMX and CSX; and on the entire θ coverage for the IMU.
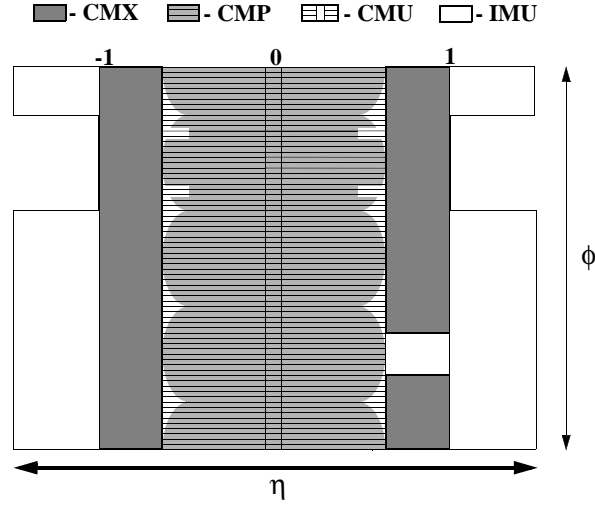
Figure 2.6: η and ϕ coverage of the Run II muon systems

2.6.1   Central muon detectors

The first muon system built at CDF, the **Central Muon Detector** (CMU) [56], is a set of 144 modules, each containing four layers of four rectangular cells. It is placed just outside the central hadronic calorimeter, whose 5.5 interaction lengths absorb more than 99% of the outgoing charged hadrons.

A second set of muon chambers, the **Central Muon Upgrade** (CMP), forms a square box around the CMU, and is shielded by an additional layer of 60 cm of steel. Due to the detector geometry, the η coverage varies with azimuth as shown in figure 2.6. The CMP consists of four layers of single-wire drift tubes, staggered by half cell per layer, and operated in proportional mode. On the outer surface of the CMP lies the **Central Scintillator Upgrade** (CSP), a layer of rectangular scintillator tiles.

Another upgrade which was begun in Run I is the **Central Muon Extension** (CMX) with the associated **Central Scintillator Extension** (CSX). It is a conical array of drift tubes, with scintillators on both sides; it extends the CMU/CMP θ coverage from 55° to 42°, except in a 30° ϕ gap which is used by the solenoid cryogenic system.

2.6.2   Intermediate muon detectors

Detection of muons in the forward region will be accomplished by the **Intermediate Muon Detectors** (IMU). This detector recycles the older Forward

Muon toroidal magnets, which will be moved closer to the interaction point (just outside the plug calorimeter PMT arrays) and will **not** be energized. The steel toroids, together with a new pair of steel rings, will then act as shielding for a new array of drift tubes and scintillator counters, placed on the outer radius of the toroids.

Like the CMX/CSX, the IMU has four staggered layers of drift tubes, and two layers of scintillator. Contrarily to the CSX, one of the scintillator layers is separated from the drift tubes by a thick layer of steel; this geometry strongly suppresses fake triggers due to hadrons.

Together with the new central tracker, the IMU will make it possible to trigger on forward muons, up to a pseudorapidity of 1.5, and to identify muons up to $\eta = 2$.

2.7 Data Acquisition and Trigger

Due to the increase in collision frequency, the DAQ and trigger systems of CDF had to be almost completely replaced. The new three-level architecture, schematized in figure 2.7, is fully capable of withstanding a 132 ns bunch separation, while keeping dead time as short as possible.
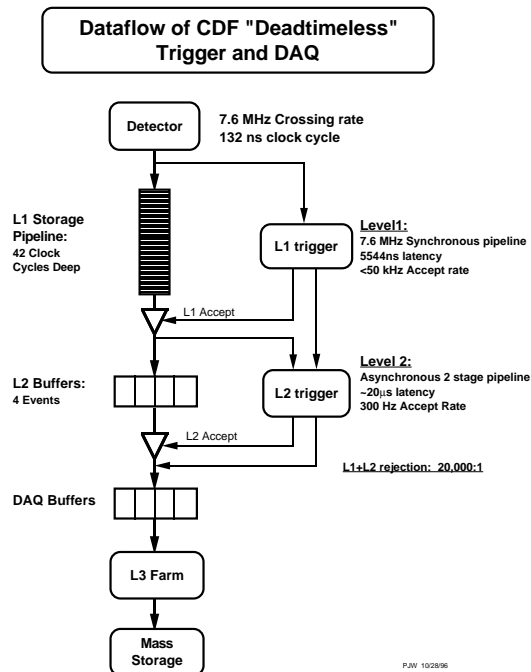
Figure 2.7: Block diagram of the CDF II Trigger

2.7.1   Level 1 trigger

The front-end electronics of all detectors is fitted with a synchronous pipeline, 42 events deep, where the entire data regarding each event is stored for 5544 ns. Meanwhile, part of the data is examined in a first layer of dedicated, synchronous, highly parallel hardware processors:

- **XFT**, the eXtremely Fast Tracker, which reconstructs tracks on the transverse plane of the COT (further described in Chapter 4), and an extrapolation unit (**XTRP**) to propagate these tracks to the calorimeters and muon chambers;

- the Calorimeter Trigger, which detects electron and photon candidates, jets, total transverse energy, and missing transverse energy;

- the Muon Trigger, which matches XTRP tracks to stubs in the muon chambers.

"Objects" from the level one trigger subsystems are combined in a flexible decision module, which takes a decision by requiring the presence of a certain number of features in the event: for example, two muon candidates with P_t above 3 GeV. Up to 64 different sets of requirements can be checked at the same time; each of these triggers can be prescaled independently of the others.

The level 1 trigger will take a decision within 4 μs, while the event's data is still in the pipeline. This makes the first trigger level truly deadtimeless. The rejection factor is expected to be about 150, thus decreasing the event rate from 7.6 MHz to about 50 kHz.

2.7.2   Level 2 trigger

Events matching the requirements of level 1 are downloaded into one of four asynchronous event buffers, and further analyzed by a second set of hardware processors. Trigger level 2 is asynchronous: events remain in the buffer until they are accepted or rejected. This can cause dead time, when all four buffers are full. In order to keep dead time at 10%, with a level 1 rate of 50 kHz, level 2 has been split in two pipelined steps of 10 μs each.

In the first phase, trigger level 2 examines event features in single detectors.

- Jets usually affect more than a single calorimetric tower. Calorimeter clustering (**L2CAL**) sums the energies collected by single towers and provides a measurement of the total jet energy.

- The calorimeter shower maximum (**XCES**) is used to reduce the rate of fake electrons and photons. It also makes it easier to match XFT tracks to their calorimetric clusters.

- The Silicon Vertex Tracker (**SVT**), which will be described in detail in Chapter 5, reconstructs tracks in the vertex detector, measuring their impact parameter d. Triggering on d will prove extremely helpful in b-quark physics.

- Data is also collected from the level 1 track and muon triggers.

During the second pipelined step, the results of the first phase are fed to a set of Alpha processors; each processor examines the event for a different set of characteristics.

Level 2 is expected to test each event for about one hundred different triggers. Requirements can range from "a single energetic lepton" to "two tracks from a secondary vertex, within a given invariant mass window". Acceptance rate will be dominated by single lepton triggers.

The level 2 accept rate will be around 300 Hz, with a rejection of about 150.

2.7.3   Level 3 trigger

After being accepted by the level 2 trigger, the entire event data is read out and loaded into a Linux PC farm, where the event is fully reconstructed in software. The level 3 reconstruction program [57] is almost fully written in C++, using object–oriented techniques; the same code will also be used, possibly with different parameters, in offline event analysis. Chapters 7 and 8 will describe the algorithms used in level 3 and offline tracking.

After an event is reconstructed, it is sent to an event counter, where its characteristics are histogrammed; if the event passes the L3 cuts, it is also permanently stored to tape.

Assuming a level 3 input rate of 300 Hz, a L3 rejection of 10, and an average event size of 250 kB, CDF II will record about $3 \cdot 10^8$ events per year, corresponding to 75 TB of data.

Chapter 3

Tracking Systems at CDF II

3.1 Introduction

Detection and tracking of charged particles is an essential part of event analysis at CDF. Trackers provide two fundamental kinds of measurement. On one side, they determine the direction and curvature of a particle's path; on the other, they delimit a narrow region where the particle might have been produced.

Charged particles moving in a uniform magnetic field, as inside the CDF tracker, have a helicoidal trajectory. By measuring the radius of curvature of the helix, one obtains the particle's transverse momentum; the longitudinal momentum is related to the helix pitch. This information can be used in several ways: as a requirement in a trigger, during particle identification, in order to calibrate the calorimeters...

To obtain a precise measurement of the helix radius and pitch, it is necessary to sample points of the trajectory which are spread on a long lever arm. Therefore, a good spectrometer requires a large tracking volume.

On the other hand, by taking a few, very accurate measurements of the track position near the primary interaction point, it is possible to narrow the region of space in which a given particle was originated. By intersecting such regions, it is possible to determine which (if any) particles were produced in a secondary vertex, trigger on their existence, and measure the mass and lifetime of short–lived particles.

Secondary vertex detection does not require a large lever arm; the most important issues are the detector's closeness to the vertex and its ability to withstand a high density of tracks.

The CDF II tracking system, shown in figure 3.1, fulfills both kinds of requirements by combining different detector elements. Momentum of tracks in

47

CDF   Tracking   Volume
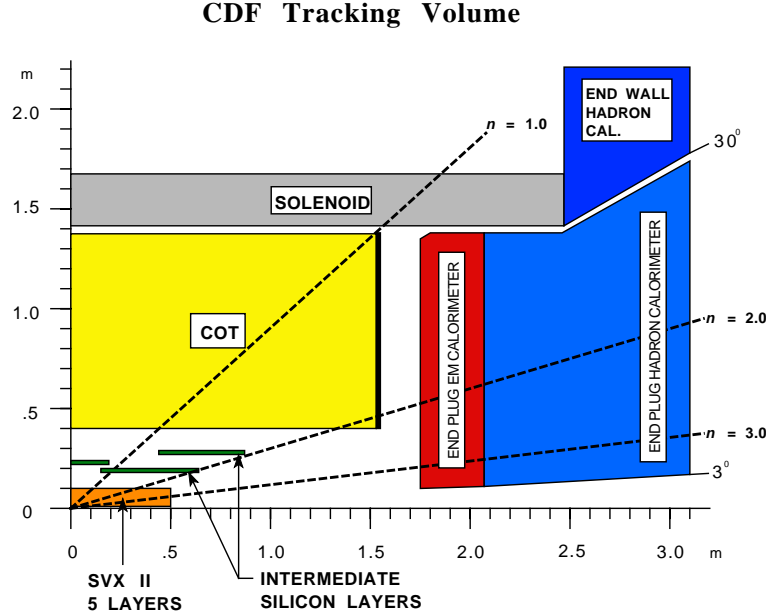
Figure 3.1: The CDF II Tracking System. Layer 00 is missing from the illustration.

the central region $(|\eta| < 1)$ is measured with an open-cell drift chamber, the **COT**, extending radially between 40 and 138 cm; several layers of micro-strip silicon wafers (**SVX II** and **Layer 00**) provides three-dimensional vertexing at radial coordinates below 10 cm; lastly, another silicon micro-strip detector (**ISL**) is used to track particles in the forward region, which is not adequately covered by the COT.

This chapter will describe in depth the structure of these detectors, after a short introduction to the conventions used at CDF to parametrize tracks.

3.2   Helix parametrization at CDF

Within an axial and almost uniform magnetic field $\vec{B}$, such as the one inside the CDF solenoid, particles of charge qe are subject to the Lorentz force

$$\vec{F} = qe\vec{v} \wedge \vec{B} \tag{3.1}$$

and move along helices of radius

$$r = \left| \frac{p_t}{qeB} \right| \tag{3.2}$$

At CDF, such helices are described with the following parameters. The term "minimum approach" refers to the point of the helix which lies closest to the detector axis, in the proximity of the origin.

$\cot \theta$ cotangent of the polar angle at minimum approach

c semicurvature[1]of the helix (inverse of diameter), with the same sign as the particle charge

z_0 z coordinate at minimum approach

d signed impact parameter: minimum distance between the helix and the detector axis. The sign of d is shown in figure 3.2.

φ_0 azimuthal direction of the track at minimum approach
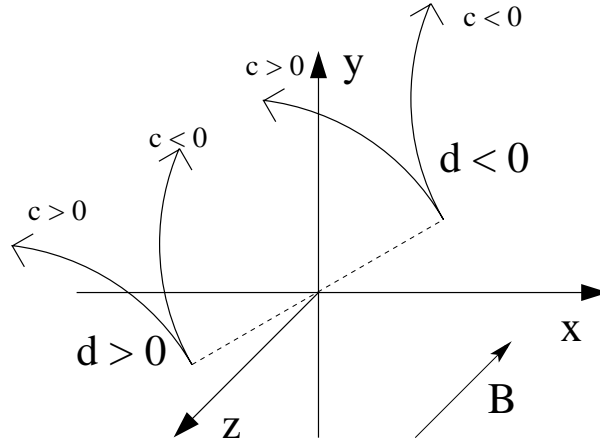
Figure 3.2: Sign of track parameters

The parametric equation of a helix in cartesian coordinates is [58]

$$
\begin{aligned}
x(\lambda) &= x_0 + q\rho \sin(q\lambda + \varphi_0) \\
y(\lambda) &= y_0 - q\rho \cos(q\lambda + \varphi_0) \\
z(\lambda) &= z_0 + \rho\lambda \cot\theta
\end{aligned}
\tag{3.3}
$$

where $\rho = \frac{1}{2qc}$ is the helix radius, $q = \pm 1$, while x_0 and y_0 are the transverse coordinates of the helix axis. Calling $S = \sqrt{x_0^2 + y_0^2}$ the distance between collider axis and helix axis, by definition

$$
d = q\left(S - \rho\right)
\tag{3.4}
$$

---

[1]Although this is not strictly correct, in CDF c is usually referred to as *curvature* instead of semicurvature.

while

$$x_0 = -qS \sin \varphi_0 \qquad (3.5)$$
$$y_0 = qS \cos \varphi_0$$

Squaring and summing the equations for x and y in 3.3, defining r as the radial coordinate of a generic helix point, and remembering that $q^2 = 1$, we obtain

$$
\begin{aligned}
x^2 &= q^2 S^2 \sin^2 \varphi_0 + q^2 \rho^2 \sin^2(q\lambda + \varphi_0) - 2q^2 \rho S \sin \varphi_0 \sin(q\lambda + \varphi_0) \\
y^2 &= q^2 S^2 \cos^2 \varphi_0 + q^2 \rho^2 \cos^2(q\lambda + \varphi_0) - 2q^2 \rho S \cos \varphi_0 \cos(q\lambda + \varphi_0) \\
r^2 &= S^2 + \rho^2 - 2\rho S \cos(\varphi_0 - (q\lambda + \varphi_0)) \\
&= \left(S^2 + \rho^2 - 2\rho S \right) + 2\rho S \left(1 - \cos q\lambda \right) \\
&= d^2 + 2\rho S \left(1 - \cos q\lambda \right)
\end{aligned}
$$

Hence

$$1 - \cos q\lambda = \frac{r^2 - d^2}{2\rho S} \qquad (3.6)$$

The projection of the helix on the transverse plane is a circumference; its equation in the polar coordinates (r, φ) is

$$r \sin(\varphi - \varphi_0) = q \left(S - \rho \cos q\lambda \right) = d + q\rho \left(1 - \cos q\lambda \right) \qquad (3.7)$$

which, making use of eq. 3.6 and remembering that $\rho = 1/2qc$, becomes

$$
\begin{aligned}
r \sin(\varphi - \varphi_0) &= d + q\rho \frac{r^2 - d^2}{2\rho(\rho + qd)} \\
&= d + q \frac{r^2 - d^2}{2(\frac{1}{2qc} + qd)} \\
&= d + q \frac{r^2 - d^2}{2(1 + 2cd)/2qc} \\
&= d + c \frac{r^2 - d^2}{1 + 2cd}
\end{aligned}
$$

This way, we reach the standard CDF equation for φ:

$$\varphi(r) = \varphi_0 + \sin^{-1} \left(\frac{cr + (1 + cd)\frac{d}{r}}{1 + 2cd} \right) \qquad (3.8)$$

Since $1 - \cos q\lambda$ is also equal to $2\sin^2(q\lambda/2)$, equation 3.6 can be also written as

$$q\lambda = 2\sin^{-1} \left(c\sqrt{\frac{r^2 - d^2}{1 + 2cd}} \right) \qquad (3.9)$$

which, when substituted in the third equation of 3.3, gives the standard equation for z:

$$z(r) = z_0 + \frac{\cot\theta}{c}\sin^{-1}\left(c\sqrt{\frac{r^2 - d^2}{1 + 2cd}}\right) \tag{3.10}$$

The purpose of the tracking system is to find tracks, and to associate to each track the best estimate of its parameters c, d, φ_0, z_0 and $\cot\theta$.

3.3 Central Outer Tracker

The Central Outer Tracker is an open-cell drift chamber, which fills the volume between radial coordinates of 40 and 138 cm, up to a $|z|$ of 155 cm. With its long lever arm, it provides an accurate measurement of track curvature, φ_0 and $\cot\theta$.

The most important design requirement for the COT is that it should be able to operate successfully at the high luminosity and collision rate of Run II, with events taking place every 132 ns. This goal was achieved by choosing a small cell size (of the order of 1 cm, four times smaller than in the COT's predecessor) and resorting to a fast gas: a 50:35:15 mixture of argon, ethane and CF_4, which, at a drift field of 2.5 kV/cm, has a drift speed of about 88 μm/ns.

3.3.1 Geometry

The COT is divided in eight superlayers, each containing twelve layers of sense wires. Axial superlayers, in which the wires are parallel to the magnetic field, alternate with superlayers in which the wires have a 3° stereo angle. Table 3.1 summarizes the most important parameters of each superlayer.

Each superlayer is divided in identical cells, whose shape is shown in figure 3.3. Cells are delimited by two field panels, made of gold-coated Mylar, and two shaper panels, which are Mylar with two field-shaping panels attached. This design bears several advantages. From the electrostatic point of view, using flat field panels instead of field wires makes the cathode surface field equal to the drift field; the drift field can therefore be higher than in a chamber using field wires. The shaper panels close the cell both mechanically and electrostatically; this allows the gap between superlayers to be small, and insures that only a single cell will become nonfunctional if a wire breaks.

The mechanical design of the COT makes use of precision-machined aluminium **endplates**, with slots in which field panels and wire planes can be inserted. The edges on which the panels are positioned have a tolerance of

Superlayer	Average r	Stereo angle	# cells	# sense wires
1	46 cm	$+3°$	168	2016
2	59 cm	0	192	2304
3	70 cm	$-3°$	240	2880
4	82 cm	0	288	3456
5	94 cm	$+3°$	336	4032
6	106 cm	0	384	4608
7	119 cm	$-3°$	432	5184
8	131 cm	0	480	5760
Axial total			1344	16128
Stereo total			1176	14112
Overall			2520	30240

Table 3.1: Central Outer Tracker geometry
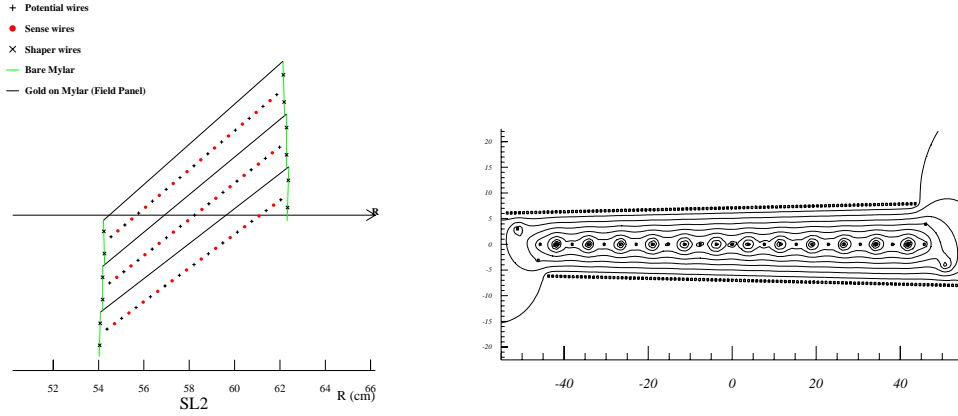
Figure 3.3: COT cell layout and equipotential surfaces

1 mil, that is 25.4 μm, in order to improve tracking resolution and reduce alignment corrections during analysis. The total load on each endplate is about 40 ton, which corresponds to a maximum deflection of about 7.6 mm; this has little effect on axial superlayers, while on stereo superlayers it leads to a small correction to the stereo angle. To minimize the negative effects of the deflection, the COT endplates have been iteratively pretensioned before any wire plane was installed.

Field panels consist of a sheet of Mylar 6.35 μm thick, with 450 Å of gold on each side, and with 300 μm stainless steel wires epoxied to the edges. The wires carry about 90% of the panel load; their shape (a slightly curved parabola) ensures the lateral tension on the Mylar is uniform.

The surface of some panels has ripples of 100 μm in amplitude; simulation shows that 300 μm ripples would cause less than 10% variation in wire gain and negligible effects on detector resolution.

Panels sag 300 μm in the longitudinal direction, and 100 μm laterally; the latter sag results in 1% variations of the drift field, but has little impact on wire gain, as effects from opposite sides of the cell cancel out. Disuniformity of longitudinal sag is more important, as a 100 μm variation between facing panels would lead to a 10% variation in gain. For this reason, field panels that differ more than 100 μm from the nominal position are discarded.

Shaper panels consist of 6.35 μm Mylar with 450 Å gold on a single side, and two 76 μm stainless steel wires on the opposite side. The gold side is grounded, while the steel wires are connected to the closest potential wires. Position of shaper panels is not critical, and may vary up to 1 mm without large adverse effects.

Wires, both sense and potential, have a tungsten core, 40 μm in diameter, plated with 450 Å of gold. They are strung at a tension of 135 g, corresponding to a nominal sag (300 μm) equal to that of the field panel, with a maximum allowed discrepancy of 100 μm. A wire support, placed at $z = 0$, ties the wires to each other, reducing the differential sag due to wire–wire electrostatic repulsion. The actual mass of the support can be varied, to make the wire sag match the field panel sag.

Potential wires are equispaced, at a distance of about 0.7 cm from each other; sense wires are placed at the midpoint between each couple of successive potential wires.

Cells are aligned at an angle of 35° from the radial direction. This tilt angle is opposite to the COT's Landau angle (which is a function of the solenoid's magnetic field, the COT drift speed, and the drift field); as a result, charges drift perpendicularly to the radial direction. Taking into account the tilt angle, the radial spacing between sense wires becomes about 0.58 cm; the overall radial lever arm of a superlayer is 6.4 cm.

	Cell area (cm^2)	Density (g/cm^3)	X_0 (g/cm^2)	$< L > / X_0$
Field panel				**0.61%**
Mylar	0.006452	1.39	39.95	0.10%
Gold	0.000091	19.32	6.44	0.12%
Steel	0.001459	7.75	13.84	0.37%
Epoxy	—	—	40.00	0.01%
Wires				**0.44%**
Tungsten	0.000324	19.30	6.76	0.42%
Gold	0.000016	19.32	6.44	0.02%
Shaper panel				**0.11%**
Mylar	0.002250	1.39	39.95	0.04%
Gold	0.000016	19.32	6.44	0.02%
Epoxy	—	—	40.00	0.00%
Steel	0.000182	7.75	13.84	0.05%
Gas	Fraction	Length (cm)		**0.54%**
Argon	50% 92.5	10.98	0,42%	
Ethane	35%	92.5	34.04	0.10%
CF$_4$	15%	92.5	64.00	0.02%
Total				**1.69%**

Table 3.2: Material in the COT active volume. Width in radiation lengths is given by formula 3.12.

Table 3.2 lists the partial radiation lengths of the component materials of the COT. By calling R the radial extent of a cell, A the cell area, ρ_i the density of each material, A_i the area occupied by that material in the cell's cross section, and X_0^i the material's radiation length, the average number of radiation lengths traversed by a radial track is

$$< L > / X_0^i = \frac{\rho_i R}{X_0^i} \frac{A_i}{A} \tag{3.11}$$

which can also be written as

$$< L > / X_0^i = \frac{\rho_i}{X_0^i} \frac{A_i}{W} \tag{3.12}$$

where W is the cell width, roughly equal to twice the maximum drift distance divided by the cosine of the cell tilt angle ($35°$): $W \simeq 2.23$ cm.

3.3.2 Operation

Voltages applied to the COT panels and wires are determined via a full-fledged electrostatics simulation program. The maximum expected cathode current, obtained by extrapolations from the Run I values, is 12.5 μA/wire, for a total of 75 μA on each side of the panel. The resistance of the gold coating is about 0.36 Ω/cm; the voltage drop along half COT is then of the order of a few millivolt, which is negligible. Due to the reduced cell size, space charge effects are going to be much smaller than in Run I.

The gas used in the COT's predecessor, during Run I, was a 50:50 mixture of argon and ethane, bubbled through ethanol. This choice would not have been appropriate for Run II. On one hand, the achievable drift speed would not have been sufficient; on the other, wire ageing in Ar-Ethane gas would happen at unacceptable rates, mostly due to alcohol mist.

The COT uses a 50:35:15 Ar-Ethane-CF4 gas mixture, bubbled through isopropyl alcohol, and filtered through a gas cleaning system. In these conditions, no significant wire ageing has been observed in irradiated test chambers, which simulated instantaneous luminosities up to 10^{33}cm^{-2}s^{-1}. As was said before, the new gas also provides a higher drift speed of 88 μm/ns at a drift field of 2.5 kV/cm.

Wire readout is performed via a custom-built ASD (amplifier, shaper and discriminator) radiation-hard 8-channel chip. ASD boards, containing three of these chips and performing readout of two adjacent cells, are placed directly on the chamber endcap. They provide continuous measurements (no external timing is needed) of the signal's leading edge and of the total collected charge; the latter quantity is encoded in the output pulse length,

and is used in offline analysis for dE/dx particle identification. In order to obtain the best compromise between dE/dx measurements and two-hit separation, the relationship between Δt and deposited charge is logarithmic $(\Delta t \simeq \log Q)$.

The signal is then carried to TDC boards, and then propagated to trigger and readout boards. These systems provide the synchronous pipeline and asynchronous buffers required by the level 1 and 2 trigger levels; they are placed outside the detector, so that they are accessible at any time.

3.3.3 Performance

The overall performance of the COT, which will be studied extensively in chapter 7, depends both on the chamber's geometry and on the quality of each channel's output. This quality can be summarized in two parameters: single-hit resolution and two-track separation.

Since the COT only measures drift time, but not drift direction, each TDC count corresponds to two possible positions of the track, placed opposite to each other with respect to the wire. It is the task of pattern recognition to associate to each hit the correct drift sign. On the other hand, the performance of pattern recognition depends on hit resolution.
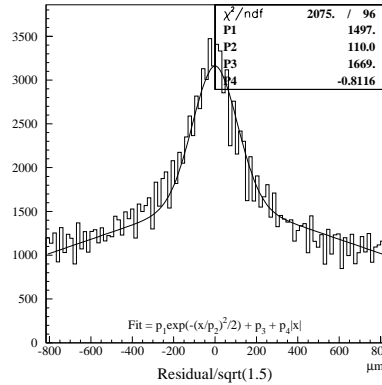
Figure 3.4: COT single-hit unbiased resolution. The x axis is rescaled, to take in account the combined error of three points.

In order to avoid dependance loops, the following procedure was used to measure the COT single-hit resolution with no bias from pattern recognition. A single COT cell, outside the magnetic field, was irradiated at a rate slow enough to avoid pile-up, and the drift times on each wire were measured. Then, for each triplet of consecutive wires, and for each of the eight possible

combinations of drift signs, the position of the two outer hits was used to predict the position of the central one. The differences between the actual and predicted positions were plotted in figure 3.4 and fitted, obtaining a single hit resolution of 110 μm.

During Run I, a similar measurement gave a result of 120 μm; the actual resolution inside the magnetic field was however 180 μm. For this reason, the overall Run II single hit resolution is expected to be 180 μm as well.

The second important parameter of COT performance is the two-track resolution, that is the ability of the chamber to reconstruct two consecutive hits on the same wire. After a hit causes the discriminator to trigger, the corresponding channel undergoes a period of dead time, until the pulse shape returns below threshold. This dead time can be converted to "dead space", multiplying it by the drift speed.

A first origin of dead space is that any amount of drifting charge, even if concentrated in a very narrow region, is transformed by the avalanche process and by the amplifier electronics into a signal of finite duration. This duration will not change much from Run I; since the Run II drift speed is higher, this process will lead to larger dead space than in Run I.
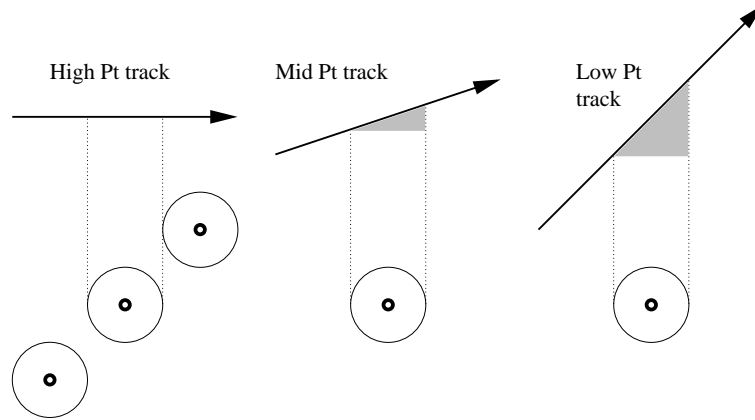
Figure 3.5: Charge drift in the COT. Electrons drift downwards; the shaded regions are the geometrical contribution to dead space.

The second, and most relevant, origin of dead space is related to the actual path followed by the particles inside the chamber. As a consequence of COT design, and as shown in figure 3.5, all the ionization charges generated by a given radial track take about the same time to drift to the avalanche region. On the other hand, charges from a low P_t or a high impact parameter track will be spread over a wide interval.

Contrarily to what happened in Run I, the COT will operate at a drift field above the v vs E saturation point; moreover, the cell boundaries will clip the dead space region for low P_t tracks. As a result, the overall average dead space will extend about 350 to 450 μm depending on the superlayer, with a 10% improvement with respect to Run I.

3.4   Silicon Tracking System

The CDF II silicon tracking system consists of eight layers of microstrips on silicon, placed in a range between 1.35 and 30 cm from the beam axis. By reconstructing the d and z_0 parameters of charged tracks, it provides accurate, three-dimensional vertexing measurements; it also extends the tracking coverage of CDF from the COT limit ($|\eta| < 1$) to $|\eta| < 2$. The tracker is divided, mechanically and for readout purposes, in three concentric subsystems: Layer 00[2], SVX II and ISL.

3.4.1   Silicon wafers

Microstrip silicon detectors are based on inversely polarized $p - n$ junctions. The bias voltage depletes the bulk material from charge carriers; if a charged particle crosses the detector, electron-hole pairs are released by ionization, and drift towards the surface. By segmenting the surface into p^+ (or n^+) strips, which are capacitively coupled[3] to conductive strips, the position of the ionizing particle can be measured along an axis orthogonal to the segments.

Several kinds of silicon wafers [59] have been used for the various parts of the tracker; neglecting some details, they can be divided in three main categories.

All of the ISL wafers, and about half of those in SVX II, are double-sided, with axial strips on the p side and **small-angle stereo** strips on the n side, which form an angle of $\pm 1.2°$ with the detector axis. The pitch varies from layer to layer, as shown in table 3.3.

To keep the construction process simple, readout is performed directly on a single edge of the wafer. This entails the cost of not being able to use the few stereo strips that do not reach the readout edge. Readout is performed by radiation-hard custom chips (SVX3), which are placed on ceramic hybrids

---

[2] The innermost layer of SVX II was originally called "Layer 0", thus the name "Layer 00". This thesis adopts the current convention of numbering SVX II layers from 1 to 5; some cited works might use the older convention.

[3] This is done to isolate the readout electronics from the diode's leakage current.

Layer	L00 0	SVX II					ISL	
		1	2	3	4	5	6	7
Axial strips:								
Implant width	8	14	15	14	14	15	22	22
Strip pitch	25	60	62	60	60	65	112	112
Readout pitch	50							
# Channels	128/256	256	384	640	768	896	512	512
Stereo strips:								
Implant width	—	20	20	15	20	15	22	22
Strip pitch	—	141	125.5	60	141	65	112	112
# Strips	—	512	576	640	512	896	512	512
2nd Al strip pitch	—	58	60	—	60	—	—	—
Multiplexing	—	4	3	—	2	—	—	—
# Channels	—	256	192	640	256	896	512	512
All pitch and width measurements are in μm.								

Table 3.3: Characteristics of silicon wafers

on top of the wafers. The chips preamplify and digitize the signal, subtract the pedestal, store the result in a pipeline during level 1 trigger operation, and transfer it outside the detector when needed.

Three layers of SVX II consist of double-sided wafers whose n-strips form an angle of 90° with the axial direction. These wafers require a more complex manufacturing technique: z metal strips are covered with an additional layer of insulator, and are read out via a superimposed set of Al axial strips. Since SVX II wafers are long and narrow, it was necessary to multiplex some non-adjacent z strips, connecting 4, 3 or 2 of them to each readout strip.

Orthogonal strips provide a very accurate measurement of z coordinates, but cause pattern-recognition ambiguities. By using both small-angle and 90° wafers, CDF II achieves good pattern recognition as well as precise 3D measurements.

Both small-angle and 90° wafers are 300 μm thick, corresponding to 0.3% radiation lengths at normal incidence. In the region covered by the hybrids, the thickness doubles to 0.6% radiation lengths.

Lastly, Layer 00, being placed extremely close to the interaction point, is built of radiation-hard, **single-sided** microstrip wafers. Kapton cables are used to carry the signal to the readout hybrids, which are placed at the z boundary of the tracker. Wafer, cable and supports add up to 0.6% radiation lengths.

It is possible to improve the position resolution, while keeping the number of readout channels constant, by reading a single strip out of N consecutive ones. The signal from disconnected ("floating") strips is divided capacitively among the two nearest connected strips, thus originating a charge cluster; the cluster centroid provides an accurate measurement of the track position. Layer 00 uses a readout pitch twice as large as the strip pitch: there is one floating strip between each pair of readout strips.

3.4.2 Geometry
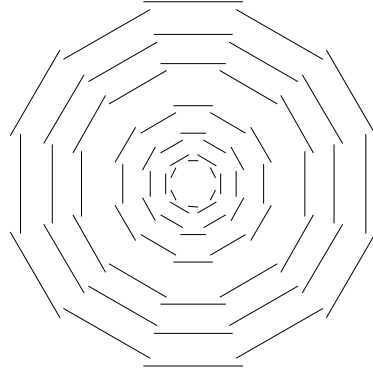
Figure 3.6: Transverse cross section of Layer 00 and SVX II. ISL is placed between two and three times the outer radius of SVX II.

Silicon wafers in L00 and SVX II are arranged on twelve azimuthal wedges, as shown in figure 3.6. In the region $1 < |\eta| < 2$, the inner layer of ISL is divided in 24 wedges, while the outer layer forms 36 wedges; in the central region ($|\eta| < 1$) 28 wedges form a single ISL layer.

Adjacent wedges partially overlap, in order to reduce gaps and increase the geometric acceptance of the detector. For this reason, and due to the requirement that the wafers should be orthogonal to the radial direction, the wafer radii are staggered. On Layer 00, due to the mechanical constraint of the beam pipe, two kinds of wafers, narrow and wide, are used; on every other layer the wafer sizes are uniform.

Table 3.4 summarizes the silicon geometry on the transverse plane, in even ("A") and odd ("B") wedges.

Each azimuthal wedge is divided, along the z axis, in six[4] readout regions, called "half-ladders"; some are shown in figure 3.7.

---

[4]In the outermost ISL layer, the two central half-ladders are missing.

	Layer	Active width (cm)	Wedge A radius (cm)	Wedge B radius (cm)	Type
Layer 00	0	0.64 (A) 1.28 (B)	1.35	1.62	Axial only
SVX II	1	1.536	2.995	2.545	90°
	2	2.3808	4.57	4.12	90°
	3	3.84	7.02	6.52	Stereo −
	4	4.608	8.72	8.22	90°
	5	5.824	10.645	10.095	Stereo +
ISL	6c	2 x 5.73	23.1	22.6	Stereo −
	6f	2 x 5.73	20.2	19.7	Stereo −
	7f	3 x 5.73	29	28.6	Stereo +

Table 3.4: Silicon tracker geometry. Layer 00 wafers have different widths in φ wedges A and B. ISL geometry is reported for the central(c) region ($|\eta| < 1$), where layer 7 is absent, and the forward(f) region. Radii are measured at the center of the wafer.
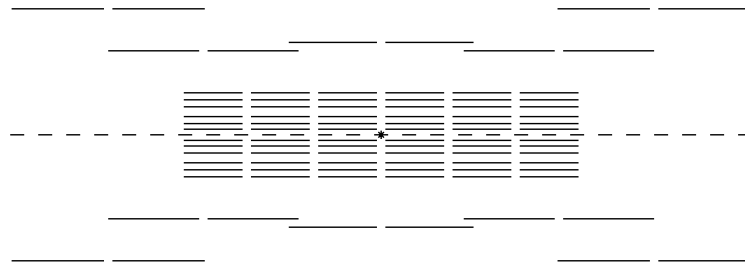
Figure 3.7: RZ cross section of the silicon tracker

Each Layer 00 and SVX II half-ladder consists of two microbonded silicon wafers. It covers a z range of about 15 cm, and is separated from adjacent half-ladders by an inactive gap of about 0.5 cm. Layer 00 and SVX II thus cover a region in z up to 46 cm from the origin; since the expected RMS spread of the interaction region along z is 29 cm, almost 90% of the primary vertices will be inside the detector.

The ISL half-ladders are longer, being formed of three microbonded wafers; they extend 22 cm each along z. In the central region ($|\eta| < 1$), a single silicon layer provides a better connection between COT tracks and SVX/L00 hits. In the forward region ($1 < |\eta| < 2$), two ISL layers help pattern recognition, and provide a long lever arm for transverse momentum reconstruction, in a region beyond the reach of the COT.

Table 3.5 summarizes the number of half-ladders and of readout channels on each layer.

Layer	L00		SVX II					ISL	
	0a	0b	1	2	3	4	5	6	7
Half-ladders	36	36	72	72	72	72	72	152	144
Axial chann./HL	128	256	256	384	640	768	896	512	512
Stereo chann./HL	—	—	256	384	640	512	896	512	512
Total axial channels	13824		211968					151552	
Total stereo channels	—		193536					151552	
Total channels	13824		405504					303104	

Table 3.5: Silicon readout channels

Operation of the vertex detector, especially in the time-critical environment of trigger levels 1 and 2, requires the detector to be well-aligned to the beam. Layer 00 and SVX II need to be coaxial to the beam with a tolerance of 100 μrad; this translates to a 15 μm misalignment between the ends of each half-ladder. Global centering of the detector is less critical: errors of up to 250 μm in the transverse position and of 1 mm along z are acceptable. ISL, being placed at greater distance from the beam, and having wider strips, is also less sensitive to alignment problems.

Since the silicon tracker is not accessible when the forward calorimeters are in place, CDF II comprises a remote alignment system: a set of piezo-electric crystals ("inchworms"), each with a step size of 50 μm, and optic fibers that can be used to measure actual wafer positions via laser pulses.

3.4.3 Performance

The overall performance of the silicon tracker, which will be studied in chapter 8, depends on the characteristics of single wafers: most importantly among these, the single hit resolution, the signal to noise ratio, and the number of dead channels.
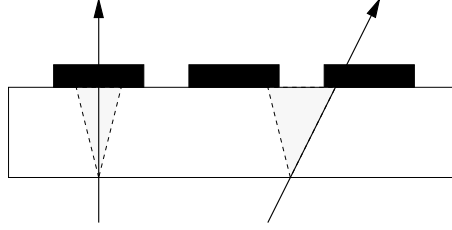
Figure 3.8: Charge deposition in a silicon detector

In SVX II and ISL, each strip is read out independently of the others. If all the charge released by a given particle is collected on a single strip, as in the case to the left in figure 3.8, the only information available to the tracker is the position of the strip. The track position is then measured with a precision of

$$\sigma = \sqrt{\int_{-p/2}^{p/2} x^2 \, dx} = p/\sqrt{12} \qquad (3.13)$$

where p is the readout pitch. This formula leads to single hit precisions of 17 μm on SVX II axial and small-angle strips, of 36 to 40 μm on 90° strips, and of 32 μm in ISL.

Better results are achieved when the deposited charge is collected by multiple strips, as exemplified in the right half of figure 3.8: the cluster shape can be analyzed to retrieve the actual position of the track. The resolution is usually better than in the single-strip case, and is limited by the signal-to-noise ratio. In the case of Layer 00, it is possible to obtain a single-hit resolution of 6 μm.

Contrarily to what happens near the wires of a gas-based drift chamber, inside a silicon detector there is no avalanche amplification of the signal; the collected charge is the same as the original ionization charge. To increase the signal it would be necessary to use thicker wafers, but this would lead to larger multiple scattering, badly affecting the overall performance of the tracker. A low noise level is therefore essential.

Noise in silicon detectors is dominated by two contributions. On one hand, the preamplifier noise depends on the total capacity of the strip, as

seen by the preamplifier itself. On the other hand, the wafer's intrinsic noise
increases as the square root of the leakage current I_{leak}. Exposure to high
doses of radiation damages the crystal lattice of silicon detectors; as a result,
the S/N ratio worsens with time.

A last factor to be considered is the presence of "dead" strips. Due to the
high redundancy of the CDF II tracking system, the loss of a single point of a
track is not an overwhelming problem in level 3 and in offline reconstruction;
it is a more important issue in the efficiency of the level 2 track trigger, SVT.

Layer 00 wafers were required to contain at most one (narrow) or two
(wide) dead strips at production; on the other layers, the majority of wafers
contains less than 2% of dead strips.

Part II

Fast Track Triggering

Chapter 4

eXtremely Fast Tracker

The eXtremely Fast Tracker [60] [61] is a synchronous, parallel, pipelined track processor: it detects high-momentum tracks on the transverse plane of the COT, quickly enough for the results to be used in the CDF II level 1 trigger. An extrapolator module (XTRP) is used to propagate the tracks found by XFT into the other parts of CDF II. Tracks matching a stub in the muon chambers provide a momentum measurement for muon candidates; those pointing at a cluster in the electromagnetic calorimeter are used in electron identification. Finally, XFT tracks are used to seed the level 2 silicon vertex tracker (SVT), which will be described in chapter 5.

By design, the XFT should be able to perform at least as well as its Run I predecessor, but at a much higher velocity. This improvement will make track information available at the first level of trigger, instead of the second (as was the case in Run I).

The requirements for XFT are the following:

- Track-finding efficiency exceeding 96% for tracks above 1.5 GeV;

- Low fake rate;

- Momentum resolution $\Delta p_t / p_t^2 < 2\%/\,\mathrm{GeV}$;

- φ_0 resolution better than 8 mrad.

With respect to Run I, such a fast, precise tracker makes it possible to loosen some acceptance cuts at trigger level 1, without exceeding the available bandwidth, and thus to collect a larger sample of interesting events.

4.1   Reconstruction Algorithm

XFT track reconstruction is performed in three steps. First of all, the TDC signal from each axial wire is digitized at a very low resolution. Digitized hits belonging to the same superlayer are then associated into line segments. Finally, compatible segments are joined to form tracks.

In the first phase, each axial hit is classified as either **prompt** (having a drift time of at most 44 ns) or **delayed** (drift longer than 44 ns). A veto window, placed between 44 and 88 ns, prevents long pulses from being interpreted as the superposition of a prompt and a delayed hit: delayed hits are reported only if they extend over the 88 ns boundary, or if there is no prompt hit on the same wire.
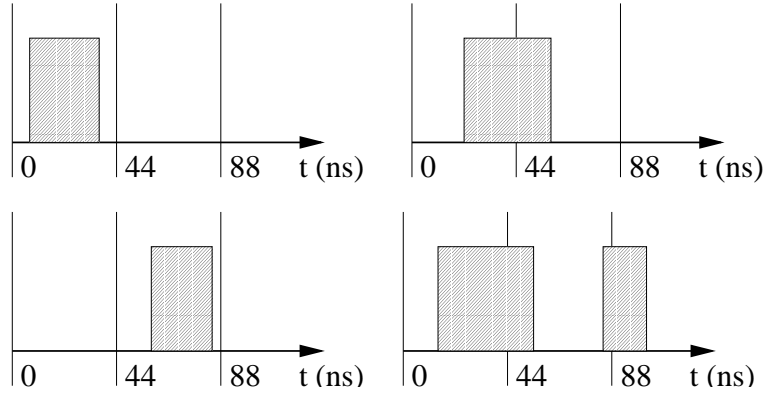
Figure 4.1: Digitization of COT hits in the XFT. The two upper hits are both reported as "prompt"; the bottom left one, as "delayed"; in the last case, both a prompt and a delayed hit are reported.

Digitized hits are then analyzed by Finder modules; each module works independently of the others, and reads the 48 wires contained in four adjacent COT cells. Each Finder contains a list of all the patterns of prompt and delayed hits that actually correspond to a line segment. By looking up the detected pattern in this list, and allowing for a certain number of wire inefficiencies, the Finder determines whether the associated cells contain a segment.

In the two inner axial superlayers, where high-p_t tracks propagate almost radially, the Finder output only describes the azimuthal position of each segment's central point, with a granularity of $1/12$ of COT cell; no information is given about the direction of the segment. In the two outer superlayers, instead, each cell is divided in only six azimuthal bins; however, the segment's direction is used to infer the charge sign (positive or negative) of the particle.
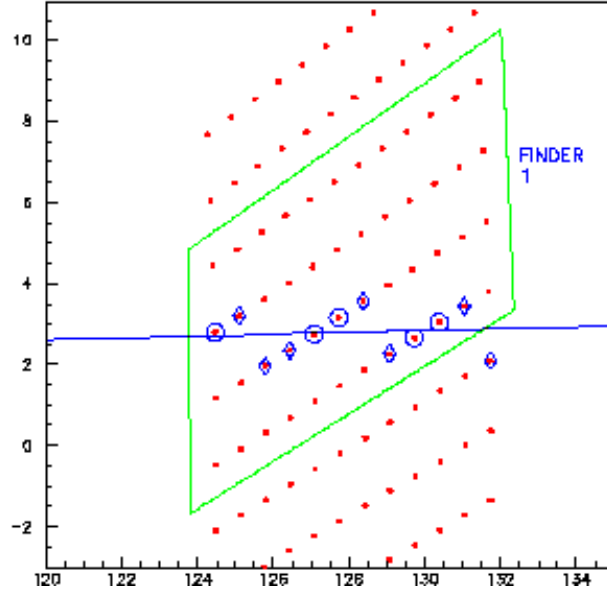
Figure 4.2: Example of XFT Finder pattern, corresponding to a high-p_t track. Prompt and delayed hits are represented as a circle and as a rhombus respectively.
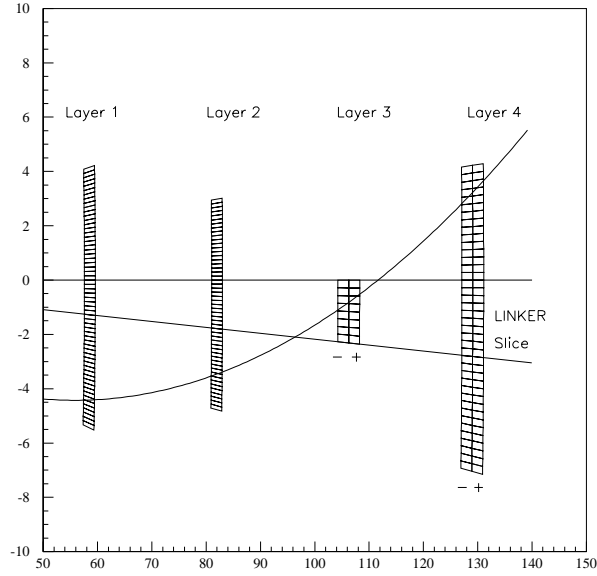
Figure 4.3: XFT Linker: example of a low-Pt track

Once the Finders complete their task, data is transferred to 288 Linker chips, each covering a 1.25° azimuthal slice of the third axial superlayer, as shown in figure 4.3. Linkers scan "roads" in the detector, looking for combinations of three or four segments that could belong to the same track. When such a combination is found, the transverse momentum and angular position of the track are reconstructed. The Linker can distinguish 48 bins in p_t for either charge sign of the particle, for a total of 96 equispaced bins in $1/p_t$; the track's φ position, measured at the center of the third axial superlayer, is determined within one of eight bins in each Linker slice.

As a final step, XFT selects the best track candidate in each slice, and reports it. Four-segment tracks are preferred to three-segment ones; if the number of segments is the same, the highest-P_t track is chosen and extrapolated to the calorimeter and muon chambers.

With some modifications to the Finder and Linker patterns, the same reconstruction algorithm could be applied to the COT stereo superlayers, and give as a result a full 3D track reconstruction. This upgrade is being considered for a future Run IIb, after 2002.

4.2   Implementation

The entire XFT is implemented in hardware, as an array of programmable logic devices: Altera FLEX 10K50 chips are used as Finders, Altera FLEX 10k70 as Linkers, and smaller auxiliary chips convert the output into the format required by CDF II's track extrapolator. Thanks to the flexibility of these devices, it will be possible to replace the search patterns at any time, for example to compensate for the presence of a "dead" wire.

The 42-event synchronous pipeline and the four asynchronous buffers, required respectively by trigger levels 1 and 2, are built on-board.

XFT is divided by data latches in several sections; at any given time, several subsequent events are being examined in parallel, each event having reached a different stage of reconstruction. The following table reports the number of 132 ns clock cycles spent in each logical phase:

Input of data from wires	1 cycle
Segment finding	1 cycle
Transfer of segments to linker	1 cycle
Segment linking	4 cycles
Output formatting	1 cycle

XFT is therefore able to find tracks on the transverse plane of the COT in 8 cycles, that is 1 μs out of the 4 μs available to trigger level 1.

4.3 Performance

The XFT group simulated the tracker's performance in software. The test sample consisted of Run I events, containing at least one muon stub, and superimposed to several minimum bias events (see figure 2.1) as expected in the higher-luminosity Run II environment. The simulation was performed both in the initial Run II situation (99% efficiency on wires, max 1 missing hit in Finder patterns) and in a much worse case (wire efficiency decreased to 92% due to accumulated damage; up to 3 missing hits allowed in Finder patterns).

The simulation's first goal was to determine the best operating parameters for the XFT: the number and width of single-hit bins, the Finder φ resolution, and the Linker P_t and φ resolution. The best performance was obtained with parameters close to those listed in the previous sections.

With the optimal XFT configuration, the overall reconstruction efficiency turned out to be almost independent of the simulation parameters, and close to the 96% design goal; transverse momentum and angular position resolution depend mostly on single wire efficiency, as shown below.

# Superimposed events	COT wire efficiency	μ rate (Hz)	$\sigma(p_t)/p_t^2$ (GeV^{-1})	$\sigma(\varphi)$ (mrad)	Reconstruction efficiency
2	99%	27.4	0.80%	2.28	96.1%
2	92%	23.1	1.03%	3.11	95.9%
10	99%	137.2	0.92%	2.74	96.7%
10	92%	151.7	1.23%	4.16	96.5%

In the expected Run II occupancy and efficiency conditions, the simulation predicts an overall fake rate below 2% for $p_t > 5$ GeV tracks. In the worst case, at extremely high luminosities, the fake rate increases to over 30%; if XFT is to operate in such conditions, it will be necessary to cross-check its results with some other detector.

It is possible to use of the XFT to implement a single muon trigger, by requiring a match between a stiff XFT track ($p_t > 10$ GeV) and a stub in the muon chambers. The simulated trigger rate is at most 150 Hz: it is a very small fraction of the level 1 budget (40 kHz), and it is already within the level 2 rate limit of 300 Hz.

XFT can also be used as a two-track trigger, with lower thresholds and without requesting a muon stub, as a prerequisite for a secondary vertex trigger. This will be the subject of chapter 5.

Chapter 5

Silicon Vertex Tracker

Early detection of secondary vertices is essential to the study of heavy quark physics in a hadronic collider. Some interesting channels, such as $B^0 \rightarrow \pi^+\pi^-$, can only be extracted from the background by requiring the presence of a secondary vertex at an early trigger level: without such a trigger, it would be impossible to collect sufficient statistics at a realistic event-recording rate,

The Silicon Vertex Tracker [62], or SVT, is a track processor designed to detect secondary vertices at the second level of trigger in CDF II. SVT reads, as its input, the list of axial COT tracks found by XFT, and the data from four axial silicon layers. It filters XFT tracks, discarding those with p_t below 2 GeV; it associates a set of silicon hits to each high-p_t XFT track, and fits the results to a circle in the transverse plane, thus determining the impact parameter.

The time allotted to the level 2 trigger is of the order of 20 μs; SVT aims at performing its task spending about 10 μs on the average event.

5.1   Architecture

SVT is divided in several subsystems, each performing a specific task.

First of all, the charge from above-threshold axial strips is digitized and read out into the processor. The detector is divided, for readout purposes, in 72 **sectors**, each containing all the half-ladders placed on various layers at the same φ and z coordinates. Readout is performed sequentially on each sector, with empty channel suppression, at a rate of 30 MHz. Since the average event will contain about 6000 active strips, the expected readout time is 3 μs.

Hit Finders, one per sector, receive the digitized pulse heights, group the active strips into clusters, and calculate the cluster centroid position. The

hits are then stored into **Hit Buffers** for future reference, and propagated to the Associative Memory units.

Pattern recognition in the SVT is performed by **Associative Memory** (AM), a massively parallel mechanism containing a long list of "roads" to be examined. A road is a coincidence between several detector elements: hits on four of the silicon layers, and XFT tracks. Upon receiving the list of hits and XFT tracks, each AM chip searches all of its roads in parallel; all of the active roads are then reported back to the Hit Buffers. Each road is entirely contained within its φ wedge; however, roads may extend between adjacent sectors along z.

A drawback of this approach is the extremely high number of roads that would be needed to match all of the possible tracks. In order to reduce the storage space needed, and bring production cost to a reasonable figure, the AM system groups hits into "superstrips", each covering a range of about 250 μm. This reduces the number of channels, but worsens the spatial resolution, increases the number of fake tracks, and may cause multiple track candidates to fall within the same road. The superstrip width was chosen as a good compromise between cost, performance, and processing time.

Another resources–performance trade-off consists in filling the Associative Memory with the most frequently active roads only. By neglecting the roads that cover a small volume of parameter space, the memory size can be reduced by about a factor of two, at the cost of a decrease in efficiency of the order of 10%.

When the Associative Memory has determined that a road might contain a track, the road's hits are retrieved from the Hit Buffer, and passed to a **Track Fitter**. If multiple hit candidates are present on the same layer, all possible combinations are checked.

Track candidates are characterized by a set of six numbers (indicated as x_i): the (full-resolution) position of the silicon hit on each of the four layers used by SVT, and the curvature and φ position of the seed XFT track. These measurements are fitted to a circumference, which is described by three parameters p_k (curvature, angular position and impact parameter). Hits belonging to a track must therefore obey three independent constraints, which can be written in the form $f(x_i) = 0$ (or $\simeq 0$, if one takes into account the finite resolution of the detector). The fitted track parameters p_k are also functions of the x_i.

Both constraints and track parameters can be linearized around some central[1] 6-tuple x_i^c, and written as scalar products that can be computed by

---

[1]Simulation studies showed that a single Taylor expansion, performed at the center of each 30° azimuthal wedge, is sufficient to achieve adequate resolution on the entire wedge.

a DSP:

$$f_j(x) \simeq f(x^c) + \sum_{i=1}^{6} \left(\frac{\partial f}{\partial x_i} \right)_{x_i^c} (x_i - x_i^c) \simeq 0 \tag{5.1}$$

$$p_k(x) \simeq p_k(x^c) + \sum_{i=1}^{6} \left(\frac{\partial p_k}{\partial x_i} \right)_{x_i^c} (x_i - x_i^c)$$

For each candidate, a χ^2 is computed from the linearized constraints f_j and from the constraint covariance matrix[2]. SVT accepts track candidates whose χ^2 is below a predetermined threshold; the linearized parameter estimates are then used in the trigger decision process.

SVT is implemented in programmable hardware. The reconstruction algorithm is hardwired; the associative memory patterns and the linear fitting constants are uploaded at the beginning of each run, and can be changed to compensate for small misalignments of the detector or of the beam.

5.2 Vertex Reconstruction and Alignment

Pinpointing the primary vertex in a b-physics interaction is a difficult task. The track multiplicity is often small; moreover, due to the high luminosity of Run II, each event is likely to contain multiple primary vertices. For these reasons, SVT does not even attempt the task. Instead, it looks for a vertex significantly displaced from the beam.

The CDF II beam spot has a transverse σ of approximately 30 μm, while b physics events often contain tracks with an impact parameter in excess of 100 μm. Therefore, neglecting the primary vertex position does not entail a major loss of performance.

Since SVT only reconstructs tracks in the transverse plane, it is of paramount importance for the detector elements to be correctly aligned to each other and to the beam. If the beam is not parallel to the axial strips, primary vertices at different z coordinates are reconstructed with different impact parameters. As a consequence, trigger performance worsens.

To avoid this, the silicon layers used by SVT are mechanically aligned to each other with extreme precision. During the first few minutes of each run, the beam position and tilt angle are measured by SVT itself. The latter is then fed back to the Tevatron and is corrected. SVT is more tolerant of

---

It is not necessary to perform different expansions for different parts of the same wedge.

[2]By choosing an appropriate set of constraints, the covariance matrix can be made equal to the identity. This decreases the number of operations, and makes SVT faster.

parallel displacements in the beam position: by changing the AM patterns
and Fitter constants, each sector can compensate such a shift.

5.3   Simulation

SVT is implemented as a series of hardware boards, containing custom chips
and large programmable logic circuits. In order to be able to check whether
a board is fully functional, it is necessary to predict its behaviour at the
level of single bits; for this reason, a full software simulation of SVT was
written, including all the aspects of track reconstruction — from hit buffers
to associative memories and track fitters.

The SVT simulator started as a standalone C program, `svtsim`, which
reads its configuration (superstrips, associative memory contents, linear fit
constants...) and event data (XFT stubs and silicon hits) from ASCII files.
`svtsim` can be interfaced to the rest of the CDF II simulation in two ways:
either by resorting to a C++ wrapper, or by using a simple translator
(`dumpSVT`) to convert event simulations from the CDF II file format into
`svtsim`'s format. The results presented in this chapter were obtained using
the latter method.

5.3.1   Data samples

The functionality of SVT and of its simulation was tested on several Monte Carlo
data samples:

- "Ideal" particles

- Noiseless single track events

- Single track events, with detector noise

All samples were propagated through the detector via the full CDF II simu-
lation, based on GEANT 3.

Ideal particle events contain a single particle, with p_t above the XFT
threshold of 2 GeV and null impact parameter; its other parameters are
randomly chosen in a region loosely containing the SVT fiducial region ($|\eta| <
1$; $z_0 = 0 \pm 29$ cm). All physical effects, such as multiple scattering, energy
loss and delta ray emission, are switched off; the particle propagates on
a perfect helix, leaving the clearest possible track in the detector. This
sample was generated mostly as a cross-check: any unexpected results in its
reconstruction would have indicated a bug in the simulation.

Noiseless single track samples resemble ideal particles, but in this case physical effects are switched on. They simulate propagation of a single muon inside CDF II, neglecting the effects of detector noise. These events were used to determine the best-case SVT resolution as a function of p_t, and the best-case efficiency for reconstruction of single tracks. 10^4 events were generated in each of 10 p_t bins: 2—3 GeV, 3—4 GeV, and eight other bins each covering an interval of 2 GeV, up to 20 GeV.

Single track with noise events are similar to the previous case, but the signal to noise ratio in the silicon detector is no longer ideal. A random amount of charge is deposited on each axial strip; the ADC threshold is chosen so that, in absence of superimposed signal, a predetermined fraction of the strips is activated by noise. When a track is present, the noise charge is summed to the charge deposited by the actual particle; as a consequence, the centroids of signal clusters may be shifted by a random amount with respect to the noiseless case. Moreover, hits caused entirely by noise can be a source of confusion during pattern recognition.

Several noise levels were simulated: 1%, which is close to the foreseen noise during actual data taking; 0.5%, 2% and 4%, to study the decrease of tracking quality with the increase of noise. Noisy events consume plenty of disk space[3]: we were only able to generate 15000 events at 1% and 2% (3000 per p_t bin, in the five most relevant bins) and 6000 events at 4% (3000 low-p_t and 3000 high-p_t).

5.3.2 dumpSVT

After being simulated with GEANT3, the events were translated into the ASCII format required by `svtsim`. The translator program consists in a series of C++ modules, most of which are part of the standard CDF II offline reconstruction code and are used to collect the required data; one module was written to perform the actual ASCII dump.

`dumpSVT` begins by reading each event from file into memory; it then reconstructs hits in the silicon, by using the clustering algorithm discussed in section 8.1; it finds axial tracks in the COT, and it finally writes hit and track data onto a human-readable file.

Silicon clustering is performed here, rather than by the simulation of the SVT hit finder, mainly because of disk space usage: "raw" strip data takes much more space than clustered hits, especially if detector noise is switched on. Moreover, the SVT simulator interface was frozen before the introduction

---

[3]The disk footprint of noisy events is about 100 kB per event, multiplied by the noise percentage. 10^4 events at 1% noise consume 1 GB.

of Layer 00; in order to make SVT use Layer 00 data without major changes in the code, the simplest solution consists in replacing the SVT hit finder with the offline one.

Two paths are available for COT tracking. One, which was used for preliminary studies, makes use of the full tracking algorithms which will be described in Chapter 7, and is able to provide a comparison between Monte Carlo and reconstructed track parameters. The other performs a bit-level simulation of XFT; its results are more realistic (during real data-taking, SVT will receive track data directly from XFT), and have a worse resolution.

In the end, dumpSVT writes out data to be loaded into SVT's hit buffers: the position of silicon hits on the layers of interest, the COT axial track stubs (curvature and φ position on the third axial superlayer), and — if the offline COT reconstruction was used — the corresponding Monte Carlo track information.

5.3.3   SVTSIM

After being translated by dumpSVT, events are loaded into svtsim, where they undergo a detailed simulation of what would happen in SVT hardware. Full-resolution hits are loaded in the hit buffers; they are grouped into superstrips and compared to associative memory contents, in order to obtain a list of active roads; the full-resolution hits from each road are fitted to a circle, and if the fit χ^2 is sufficiently small the hits are declared part of the same track.

The output of svtsim consists in a series of histograms and ntuples describing the set of reconstructed tracks. For each track, svtsim stores the event number, the fit parameters, the wedge and half-ladder, the positions of the track's hits, and, if available, the original Monte Carlo track parameters.

5.4   Performance with the baseline detector

SVT performance can be characterized at two different levels: track level and event level. At track level, SVT must be able to recognize the passage of a particle with above-threshold momentum (efficiency), to reconstruct the impact parameter accurately in most cases (resolution), and to make large errors rarely (distribution tails). At event level, instead, SVT must strike a good compromise between acceptance of signal and rejection of background.

In the baseline design of CDF II, SVT uses axial hits from the four inner layers of SVX II. This section describes SVT performance in the baseline configuration, at track level; event-level studies are in progress.

5.4.1 Track resolution

SVT **resolution** was measured first of all on the noiseless single track sample, using the full COT reconstruction; this provides a best case estimate, to be compared to the outcome of more realistic samples.

Even and odd azimuthal wedges have different geometries, as shown in figure 3.6 and in table 3.4: SVX wafers in "B" wedges are closer to the beam than the corresponding "A"–wedge wafers. Therefore, "B" wedges provide a slightly better impact parameter resolution[4]. For this reason, resolution measurements were repeated separately for "A" and "B" wedges.

Reconstructed tracks were binned according to their transverse momentum; the impact parameters within each bin were histogrammed and fitted to a gaussian, whose σ (with its error) was taken as an estimate of SVT resolution. Figure 5.1 shows the results: the resolution at low p_t (2 to 3 GeV) ranges between 26 μm (wedge A) and 23 μm (wedge B), and it improves respectively to 14 and 12 μm for high transverse momentum.

As a cross-check, resolution was also computed on the ideal particle sample, where multiple scattering and other physical effects are switched off; the results were independent of p_t and matched the asymptotic resolution values of figure 5.1, as expected.

The same procedure was repeated on the noisy single-track samples, to study the effects of noise on resolution. As can be seen in figure 5.2, the presence of even a small amount of noise has a large effect on asymptotic resolution, worsening it by about 20%; further increases of the noise level, however, cause smaller increases in σ_d. This is consistent with a "random walk" model, in which each noise ADC count pushes the fit parameters in a random direction: the expected overall effect is proportional to the square root of the number of steps. At low p_t, multiple scattering already smears the track considerably, and the effects of added noise have a slower onset.

Lastly, to evaluate the effects of the XFT simulator on resolution, the simulation was repeated replacing XFT tracks to the fully-reconstructed tracks. The effects of XFT on impact parameter resolution proved to be small, and are exemplified in figure 5.3.

5.4.2 Efficiency and fake rate

Due to several reasons, SVT is not capable of reconstructing every track that crosses CDF II's central tracker. The purpose of this section is to estimate how often and why this happens.

---

[4]The situation is reversed when Layer 00 is taken in consideration: on Layer 00, the "A" wedge is closest to the beam, and provides the best resolution.
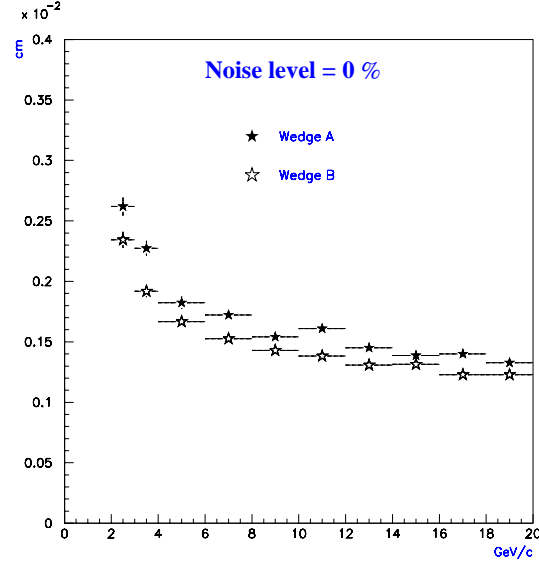
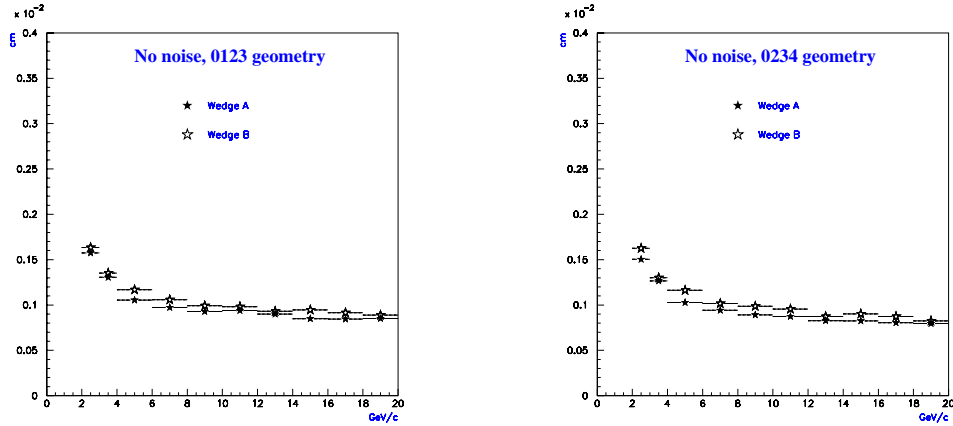Figure 5.1: Baseline SVT d resolution as a function of p_t, on a single-track, noiseless sample using full COT reconstruction.
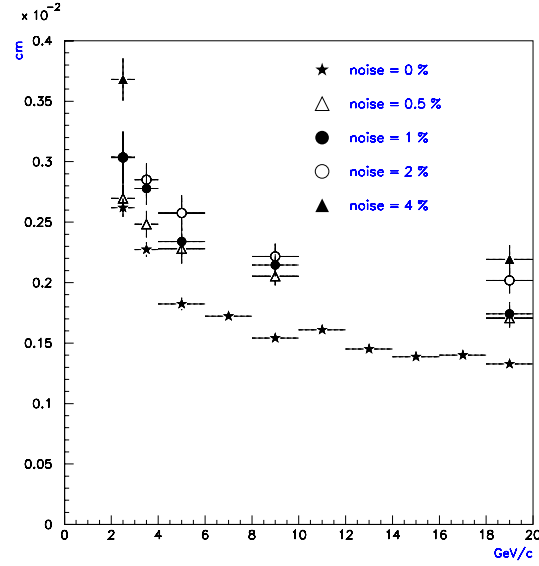
Figure 5.2: Baseline SVT d resolution as a function of p_t, on single-track samples with different noise levels. Only wedge A data is shown.
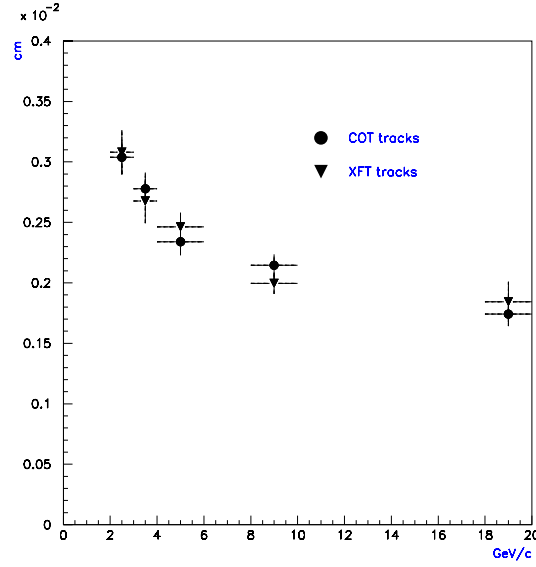
Figure 5.3: Baseline SVT d resolution as a function of p_t: comparison between usage of full COT tracking and XFT simulation (1% noise level, wedge A)

On one hand, there are some "hard" requirements to track reconstruction. SVT defines a track as a coincidence of an XFT track and of four silicon hits, placed in the same wedge and in the same sector (or in adjacent sectors). If any of these are missing (because of an XFT inefficiency, or because the track slipped through a crack in the detector on some layer, or even because the track's curvature caused it to pass from one azimuthal wedge to another), the track cannot be reconstructed. This is an intrinsic limitation of SVT.

On the other hand, as was anticipated earlier, the Associative Memories do not contain every possible pattern of hits, but only the most commonly encountered patterns. SVT might therefore be unable to reconstruct a track because its pattern was not listed. Contrarily to the previous case, these inefficiencies are due to a trade-off between performance, time spent in reconstruction, and cost.

In order to estimate the performance of SVT, the first kind of inefficiencies is not extremely relevant. First of all, they are unavoidable inefficiencies; it is impossible to "tune" the tracker to improve the situation. Secondly, the fraction of "findable" tracks on a given sample depends heavily on the kinematic characteristics of the sample itself: as shown in figure 5.4, large-η tracks have a higher chance of missing at least one layer.

It is more interesting to estimate the ratio between found and findable
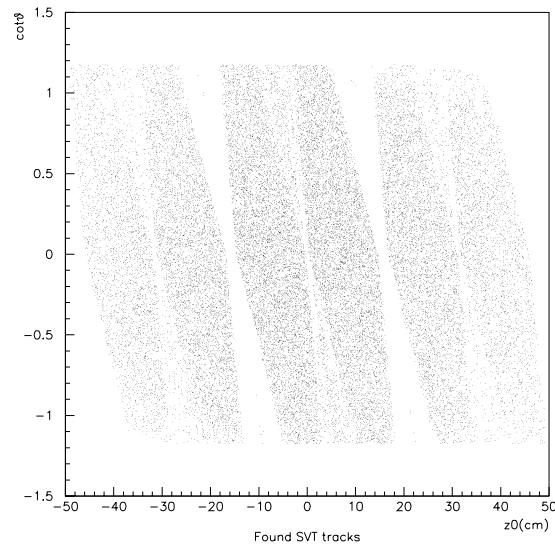
Figure 5.4: Acceptance region of SVT in terms of z_0 and $\cot \theta$. The six hexagonal regions correspond to tracks completely contained within a single sector; the smaller (and less populated) triangular regions, to tracks crossing between adjacent sectors. The initial sample had a flat distribution in η (up to 1) and a gaussian spread of 29 cm in z_0.

tracks, as a function of noise level. To achieve this goal, the first step consisted in determining the number of findable tracks on the single-track sample.

In the case of noiseless tracks, the situation is clear. All the hits belong to a "real" track; the fiducial number of findable tracks that are completely contained within one sector corresponds exactly to the number of sectors that contain at least one hit per layer. These were counted by means of a simple Perl script. Tracks that cross the boundary between adjacent sectors were neglected, both here and in the count of found tracks.

Out of a sample of 10^5 tracks, 69947 turned out to be "single-sector findable" after use of the full COT tracking algorithm; 65457 fiducial tracks were detected in the same sample, when the XFT simulator was used.

Counting the number of findable tracks on a noisy sample is more problematic. Due to the presence of noise hits, it is not possible any longer to merely count sectors: even if a sector contains one hit on each layer, most of those hits might not be related to any track. In order to estimate the efficiency denominator on noisy samples, therefore, the noiseless results were rescaled with the sample size, as shown in table 5.1. This was done under the assumption that the presence of noise hits would seldom "recover" a good track that would otherwise be lost due to detector cracks.

Noise level	Sample size	# Fiducial (COT)	# Fiducial (XFT)
0	100000	69947	65457
0.5	15000	10492	9819
1	15000	10492	9819
2	15000	10492	9819
4	6000	4197	3927

Table 5.1: Number of SVT fiducial single-track events

The following step consisted in counting the events where a track had actually been reconstructed by SVT inside a single sector. Events were counted, rather than tracks: it is possible for SVT to associate the same XFT track to multiple sets of hits in the silicon. Efficiency was defined as the ratio between the number of events where the track was found (with an impact parameter error of less than 100 μm) and the number of events with a findable track. Tracks with a larger error were computed separately, and used to estimate the fake rate[5]. The same event could contribute both to efficiency and fake

---

[5] A threshold of 100 μm was used, because the impact parameter cuts actually used in the level 2 trigger are usually of the order of 100 μm.

rate, if the same XFT track had been associated both to a "good" and to a
"bad" set of silicon hits. Table 5.2 reports the results on single-track events,
as a function of silicon noise.

COT type	Noise %	Reconstructed #	Reconstructed %	Large-error #	Large-error %	Errors/Reco %
	0	61981	89%	162	0.2%	0.3%
	0.5	9206	88%	204	1.9%	2.2%
Full	1	9201	88%	324	3.1%	3.5%
	2	9190	88%	528	5.0%	5.7%
	4	3645	87%	468	11%	13%
	0	56605	86%	145	0.2%	0.3%
XFT	0.5	8326	85%	183	1.9%	2.2%
	1	8332	85%	279	2.8%	3.3%

Table 5.2: Baseline SVT: single track intrinsic efficiency and fake rate as a
function of noise level. High-noise samples were not reconstructed with the
XFT simulator due to computing resource limitations.

As can be seen from the table, the ratio between "good" and "bad"
reconstructed tracks is independent of the actual method used to find tracks
in the COT: using the full COT tracking algorithm or the XFT simulator only
leads to rescaling the overall efficiency and fake rate by the same amount. The
efficiency is almost independent of noise, even for high levels of occupancy;
the fake rate is roughly proportional to the number of noise hits, which leads
to the hypothesis that most fake tracks consist of three "real" hits associated
to a single fake hit.

5.5   Usage of Layer 00 in SVT

The addition of Layer 00 to CDF II prompted a study on its possible use
inside SVT: with its precise measurement and its closeness to the interaction
point, Layer 00 can greatly improve impact parameter resolution.

Two geometries were taken in consideration. The first, labeled "0123",
involves replacing the outermost layer of baseline SVT with Layer 00; SVT
would then use L00 and the three innermost SVX II layers. From a geometric
point of view, this is undoubtedly the best choice: impact parameter is mea-
sured using the detectors closest to the interaction point, and the amount of
material crossed by particles before they leave the trigger region is minimal.

The second geometry, instead, is called "0234": it uses Layer 00 to re-place the *innermost* layer of baseline SVT, that is the innermost SVX II layer, which — according to some studies — could suffer from degraded per-formance due to radiation damage. Layer 00, being radiation-hard, would be a longer-lived replacement. The main disadvantage of this configuration is the presence of an unused silicon layer inside SVT. This makes pattern recog-nition more difficult in two ways: it introduces additional multiple scattering within the SVT tracking volume, and it enlarges the tracker lever arm, thus amplifying the effects of multiple scattering. As a consequence, geometry 0234 is expected to provide a worse efficiency and a higher fake rate; reso-lution on correctly reconstructed tracks, being dominated by the Layer 00 point, should be similar to the 0123 case.

The performance of SVT in the alternative geometries was measured in the same way as in the baseline configuration. Figures 5.5 and 5.6 show the impact parameter resolutions; figure 5.7 compares the results for the three geometries. As far as resolution is concerned, the addition of Layer 00 to SVT is a clear advantage with respect to the baseline configuration.

Figure 5.5: SVT d resolution as a function of p_t, on a single-track, noiseless sample using full COT reconstruction. Geometry 0123 is shown to the left, 0234 to the right.

Efficiencies and fake rates are shown in tables 5.3 to 5.6. The consid-erations that were stated earlier for the baseline geometry are still valid: efficiency is almost independent of noise, the fake rate increases almost lin-early with noise, and the wrong-to-good track ratio is independent of which algorithm was used to simulate COT track reconstruction.
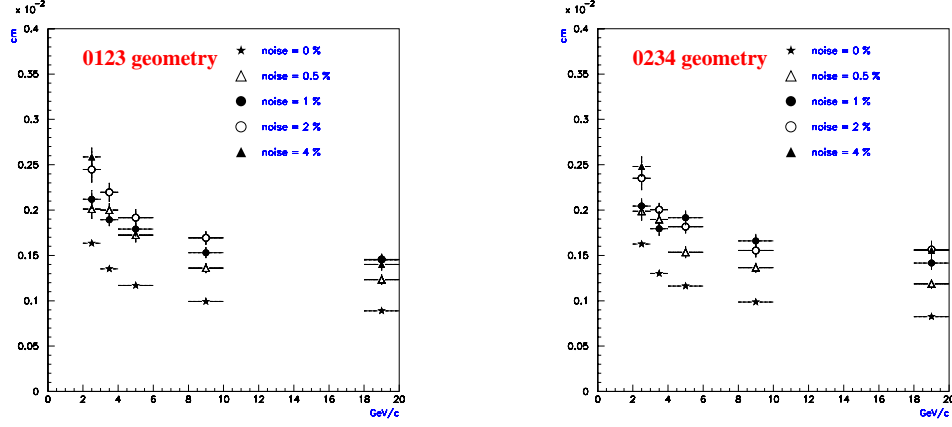
Figure 5.6: Baseline SVT d resolution as a function of p_t, on single-track samples with different noise levels. Only wedge B data is shown.
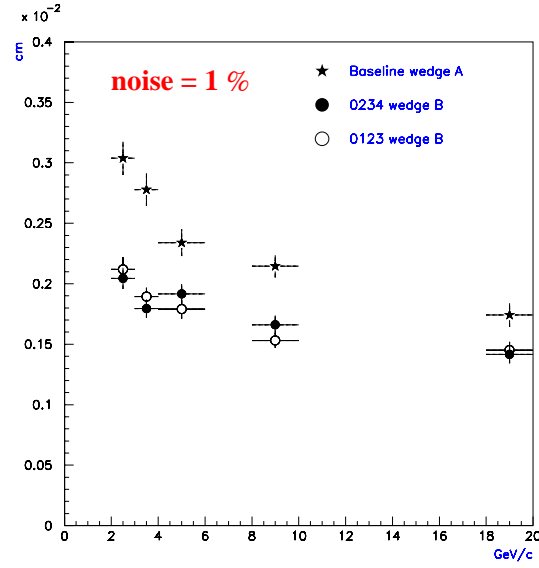
Figure 5.7: Comparison of SVT impact parameter resolution on a single-track sample, with 1% noise, on the three SVT geometries. Only data on the worst wedge (A for baseline, B for Layer 00) is shown.

Noise %	Events	Fiducial events			
		0123		0234	
		Full COT	XFT	Full COT	XFT
0	100000	72085	67453	70290	65654
0.5	15000	10813	10118	10543	9848
1	15000	10813	10118	10543	9848
2	15000	10813	10118	10543	9848
4	6000	4325	4047	4217	3939

Table 5.3: Estimated fiducial single-track events for alternate SVT geometries

COT type	Noise %	Reconstructed #	%	Large-error #	%	Errors/Reco %
	0	65712	91%	15	< 0.1%	< 0.1%
	0.5	9814	91%	134	1.2%	1.4%
Full	1	9821	91%	230	2.1%	2.3%
	2	9764	90%	380	3.5%	3.9%
	4	3868	89%	372	8.6%	9.6%
	0	59947	89%	16	< 0.1%	< 0.1%
XFT	0.5	8857	88%	111	1.1%	1.3%
	1	8856	88%	201	2.0%	2.3%

Table 5.4: SVT, "0123" geometry: single track intrinsic efficiency and fake rate as a function of noise level.

COT type	Noise %	Reconstructed #	%	Large-error #	%	Errors/Reco %
	0	57890	82%	101	0.1%	0.2%
	0.5	8607	82%	415	3.9%	4.8%
Full	1	8635	82%	572	5.4%	6.6%
	2	8592	81%	791	7.5%	9.2%
	4	3409	81%	626	15%	18%
	0	52552	80%	90	0.1%	0.2%
XFT	0.5	7767	79%	369	3.7%	4.8%
	1	7777	79%	507	5.1%	6.5%

Table 5.5: SVT, "0234" geometry: single track intrinsic efficiency and fake rate as a function of noise level.

Noise %	Geometry					
	Baseline		0123		0234	
	Reco	Errors/Reco	Reco	Errors/Reco	Reco	Errors/Reco
0	89%	0.3%	91%	< 0.1%	82%	0.2%
0.5	88%	2.2%	91%	1.4%	82%	4.8%
1	88%	3.5%	91%	2.3%	82%	6.6%
2	88%	5.7%	90%	3.9%	81%	9.2%
4	88%	13%	89%	9.6%	81%	18%

Table 5.6: Summary of SVT efficiency and fake rate

The impact of SVT geometry on results is evident. As expected, the "0123" geometry leads to a performance improvement. Efficiency undergoes a slight increase; more importantly, the single-track fake rate at the expected 1% level of noise decreases from 3.5% to 2.3%. Assuming that SVT will be used to implement a trigger which requires two tracks with $|d| > 100$ μm, this improvement should reduce the trigger fake rate by a factor of two.

The opposite is true for the "0234" geometry: the presence of a gap between the SVT layers causes a performance drop. Single-track efficiency is noticeably lower than in the other cases. Due to the longer lever arm of the geometry, a higher number of roads would be needed to obtain the same efficiency as before; since the associative memory size must be kept constant, the efficiency decrease is unavoidable. Moreover, the geometry gap between the SVX layers and Layer 00 makes pattern recognition more difficult: wrong hits are associated more frequently to the SVT track. This results in an increase of the single-track fake fraction to 6.6%; the corresponding trigger fake rate is about three times larger than in the baseline geometry.

Work is currently in progress in order to obtain a more direct estimate of trigger-level SVT performance. Procedures similar to the ones described earlier are being applied to **physical background** samples, where a uniform noise is superimposed to an uninteresting physical event, and to **signal** events. Previous studies [63], performed in absence of physical effects (such as multiple scattering) on the "0234" geometry, agree qualitatively with the single-track conclusions above.

Another work in progress involves evaluation of the "2345" SVT geometry, in the hypothesis of replacing a radiation-damaged innermost SVX II layer with the outermost layer, instead of Layer 00. While this approach would avoid gaps inside SVT, and would keep the lever arm short, it would also lead to a poor resolution on impact parameter.

Part III

Full Track Reconstruction

Chapter 6

Data samples and analysis tools

To evaluate the performance of a tracking system, it is necessary to define a benchmarking standard, to apply the system to some test cases, and to compare the actual outcome to the expected results. This chapter describes the tools that were used in these tasks.

6.1 Monte Carlo samples

An essential tool in tuning and testing a tracking system consists in a set of Monte Carlo–simulated events. Contrarily to real data, MC events are available before the detector is fully built; beside the detector output, they contain additional information about what "really" happened (for example, which particle originated which hits, and the real geometry of the event).

6.1.1 Simulation process

The CDF II simulation program, `cdf2sim`, performs a full-scale simulation of the collision process and of the detector response.

First of all, an event generator is used to obtain an initial set of particles. `cdf2sim` contains several interchangeable generators, ranging from `FAKE_EVENT`, which randomly "shoots" particles of any required type in the detector, to `Pythia` and `Herwig`, which perform a fairly realistic simulation of $p\bar{p}$ collisions and heavy particle decay.

By default, an event's primary vertex is placed in the origin, and happens at $t = 0$; it is possible to smear its position in space and in time, in order to take the beam size into account.

Each particle then propagates into the detector, interacts with the material, releases charge in the tracker and possibly decays or is absorbed; the

detector elements are read out and digitized. These processes are simulated in GEANT 3–based code.

The final output is organized in "banks", in the same format as the real detector's readout: this simplifies the task of testing reconstruction software, as the same code can be used both on simulated and on real data without modifications. Each bank contains information from a detector subsystem, such as the COT; Monte Carlo events also contain "truth" banks (obviously not available in real data) which list all the simulated particles, their properties, and the effects they had on the detector.

6.1.2   Generated samples

Two kinds of simulated events were used as a benchmark for the software tracking system of CDF II: single track events, and $t\bar{t}$ events.

Single track events are the simplest events that can be simulated in the detector. Containing a single observable particle, they do not suffer from several problems that degrade performance on more complicated events. They represent therefore an ideal sample, on which the tracking algorithms should almost always succeed.

Most of the interesting tracks in a real event originate at a very small distance from the detector axis, and therefore have a small impact parameter. As will be explained in chapter 7, this feature can be used as an additional constraint during reconstruction: once the majority of low-d tracks have been found and their hits have been marked as "used", the beam constraint can be lifted and the search repeated.

For this reason, two sets of single-track events are needed: one set of tracks with no impact parameter, and another set in which d varies randomly in a sufficiently wide range (at least about 1 cm, the radius of the beam pipe).[1]

Optimal tracking on single track events is only half of the problem: the tracker must also be able to operate correctly in a high occupancy environment, rejecting spurious hits and keeping the fake track rate low.

The optimal benchmark for these requirements would be a set of typical b and electroweak physics events, with an appropriate number of superimposed minimum bias events (as discussed in section 2.2) and an adequate level of detector noise. For historical reasons[2], the CDF II tracking group chose instead a sample of $t\bar{t}$ events, with no noise and no superimposed events;

---

[1] Since `FAKE_EVENT` cannot set the impact parameter directly, the desired outcome was achieved by smearing the transverse position of the primary vertex by a larger amount, and choosing the track direction randomly.

[2] Adequate characterization of detector noise has not been available until recently; simulation of superimposed minimum bias events still suffers from some problems.

the high track multiplicity compensates for the lack of harder-to-simulate features.

6.2 Best case analysis

The fundamental task of a tracking system consists in grouping detector hits into tracks; the tracks can then be fitted, extrapolated, intersected, and used in other aspects of event analysis. The best results are obviously obtained when the pattern recognition step is performed without any error: every hit is assigned to the track that actually caused it.

In order to put an upper limit to the quality of tracking at CDF II, and to obtain a term of reference for the performance of actual tracking algorithms, I developed two analysis tools to be used on simulated events. The first is a "perfect tracker": a tracking module that reads hit parentage from Monte Carlo data, and uses it to simulate best-case pattern finding. The second is an "ideal track fitter", which determines a perfect track's parameters by applying the standard CDF fitting procedure to the track's "clean" hits, discarding those hits whose position has been shifted due to hit overlap.

6.2.1 Perfect tracking

Perfect tracking begins by counting the number of particles in the event, and creating an array of empty tracks of the same size. This choice is not efficient from a memory point of view: it would have been possible to create a track only for those particles that actually left some hits in the detector, and not for neutral particles or short-lived particles. On the other hand, by using an array, accessing the hits of a particle given the particle's number is an extremely fast operation; besides, a large fraction of particles have at least one hit.

The code then scans COT and silicon hits, and retrieves the number of the track (or tracks) that contributed to the hit. At this point, the algorithm tries to ensure that the hit actually belongs to the interesting part of each contributing particle's track, and not, for example, to the ingoing leg of a low-p_t looping particle. This is done in different ways, according to the hit type.

In the case of COT hits, the hit's radial coordinate is inserted in the helix equations 3.8 and 3.10 together with the particle's MC parameters, to obtain an estimate of the φ and z position of the track's outgoing leg at that value of r. The estimated z coordinate is then used to determine the hit's

φ position; this is a small correction (due to wire sag) for axial wires, but is more relevant for stereo wires. Lastly, the track and wire φ coordinates are compared to each other; if they match within a configurable window[3] the hit is accepted.

Axial and small-angle stereo silicon hits use a simpler procedure: the track is intersected with the plane of the hit's wafer. In order to be accepted, the hit is required to lie within a few centimeters of the intersection; this loose cut only removes hits from inbound legs of loopers.

On the other hand, 90° silicon hits require more attention, since they are multiplexed. For each hit, each contributing track is extrapolated and intersected with the wafer's plane; the multiplexed hits are then matched to the particle's path, and the one closest to the intersection is attached to the track. As in the previous case, the hit is also required to lie close enough to the Monte Carlo track.

6.2.2 Ideal track fitting

Hits found by perfect tracking may be of varied quality: some are generated by a single parent particle, others mix information coming from two or more tracks. In the latter case, the hit seldom matches both tracks well. In the COT, only the hit corresponding to the leading edge is reported. Other tracks contributing to the same hit are present in the parents list, but the track position is often several hundred micron away from the hit. In the silicon, instead, clusters from nearby tracks may merge, making the detector report a single hit in an intermediate position; conversely, a single particle may produce multiple (incompatible) hits on the same wafer, due to a broken cluster produced by noise and/or delta ray emission.

In order to determine the ideal fit of a track, therefore, it is necessary to find out which of the track's hits can be trusted and which cannot. To keep the procedure simple and under control, this was done in the following way:

- In the COT, hits were used to fit only the track corresponding to the leading edge of each TDC pulse. Other contributing tracks were fitted without this (obscured) hit.

- In the silicon, shared and split hits were discarded before the fit.

Surviving hits were then passed to the standard CDF II track fitters; the resulting estimates of the track parameters were taken as terms of reference for the actual track-finding algorithms.

---

[3]By default, a rather loose cut of 100 mrad is used.

6.3 Evaluation tools

At the level of single tracks, the performance of tracking algorithms can be studied from two points of view. On one hand, it is possible to compare the set of findable hits, as determined in perfect tracking, to the set of found hits. This approach puts an emphasis on the "real" parentage of hits; it is most appropriate in the silicon tracker, where acceptance of a single wrong hit can have major effects on the reconstructed track's parameters — and hence, possibly, on event analysis.

Alternatively, the ideal fit parameters can be compared to the real fit ones. This fit-based approach is more adequate in the COT, where — due to the high number of layers — the chance of picking up wrong hits is large; on the other hand, a wrong hit in the right place does not influence fit results much. In silicon, fit-based efficiency is expected to give similar results to hit-based efficiency.

6.3.1 Hit classification

When determining hit-based performance of a tracker, tracks are divided into groups based on the number and quality of their hits.

For each Monte Carlo particle, hits are first of all categorized according to Monte Carlo information. Some hits clearly belong to the particle's track: they caused a leading edge on a COT wire, or they formed a silicon cluster which is not shared with other tracks. These are called **clean** hits; not being able to associate one of them to its track should be counted as an inefficiency. COT hits which have been obscured, or silicon hits with a shared/split cluster, are called **dirty** hits. Since their position is not guaranteed to match the track they belong to, they are not required to be found by tracking algorithms; on the other hand, finding such hits is not a mistake either. The union of clean and dirty hits makes the set of **available** hits; all other hits are not related to the current track.

The particle's Monte Carlo hits are then compared to the track finder's output, that is to the list of found hits. Table 6.1 lists the nomenclature for the various categories.

In silicon detectors, due to the overlap region between azimuthal wedges, it is possible for a particle to have more than a single clean hit on a given layer (provided the wafers are different). On the other hand, it is also possible for a track to intersect a layer in the crack between two adjacent wafers, leaving no detectable hit. The quality of a track with two hits on one layer and no hits on another layer is different from that of a track with two hits on two different layers. For this reason, it is useful to extend the hit categories to

| MC status | Tracking status | |
	Found	Not found
Clean	Good	Missed
Dirty	Recovered	Overlooked
Other	Wrong	—

Table 6.1: Categories of hits in track finding

silicon layers. A layer is:

- clean, if it has at least one clean hit;

- available, if it has at least one available hit;

- found, if it has at least one found hit;

- wrong, if it has at least one wrong hit;

- good, if it has at least one good hit and no wrong ones.

Figures 6.1 and 6.2 show the relative abundance of clean and available COT hits and silicon layers per track, for single track and $t\bar{t}$ events. As expected, the situation worsens with the increase of occupancy.

Once a set of reconstructed tracks is available, it is possible to define which tracks have been correctly reconstructed in terms of the number of hits in the categories above: for example, one can require a "good" track to have zero wrong hits and at most N missed hits.

6.3.2   Fit-based classification

The hit classification method presented above suffers from some drawbacks. In some cases, particles give origin to "clean" hits that shouldn't be associated to the corresponding track. This can happen when some physical effect (multiple scattering, delta ray emission...) causes the actual path of the particle to deviate considerably from an ideal helix. On the other hand, accepting some "wrong" hits that happen, by chance, to lie quite close to the real trajectory does not represent a problem.

An alternative method consists in comparing not the individual hits, but the fit parameters of the ideal and the real track. Whatever its hit content, a found track is defined "good" if it matches the ideal fit within a given error.
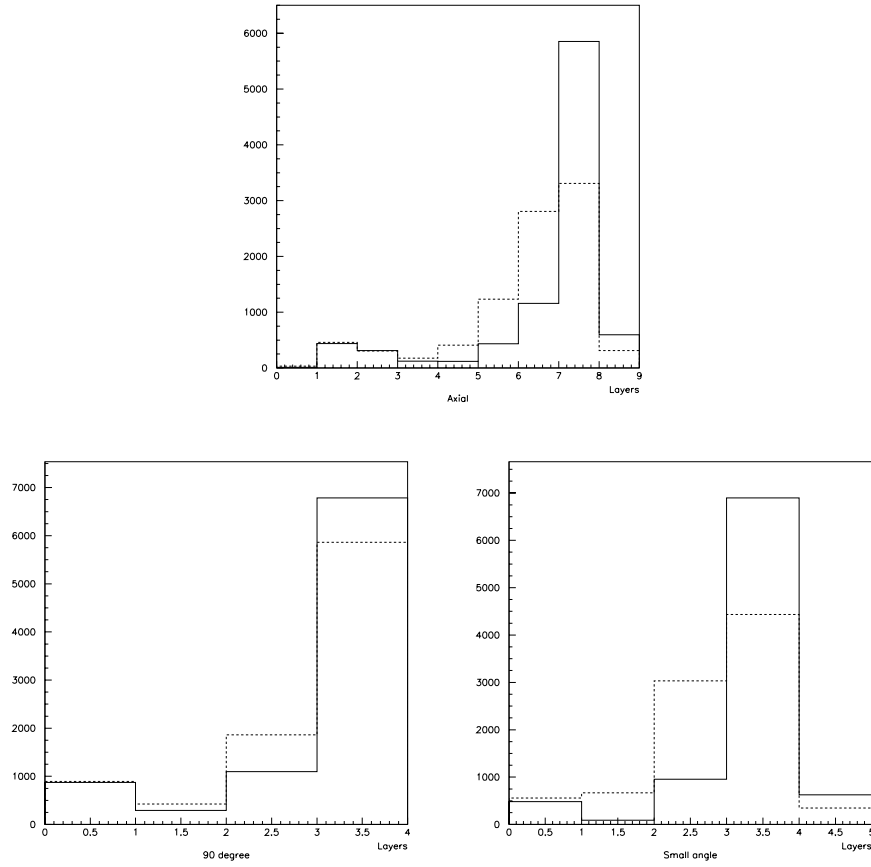
Figure 6.1: Clean (dashed line) and available (solid line) silicon layers per track, in single track events, divided by type: axial (above), 90° (left) and small-angle stereo (right). In single track events, all COT hits are clean by definition, and most tracks have a maximal number of available COT hits.
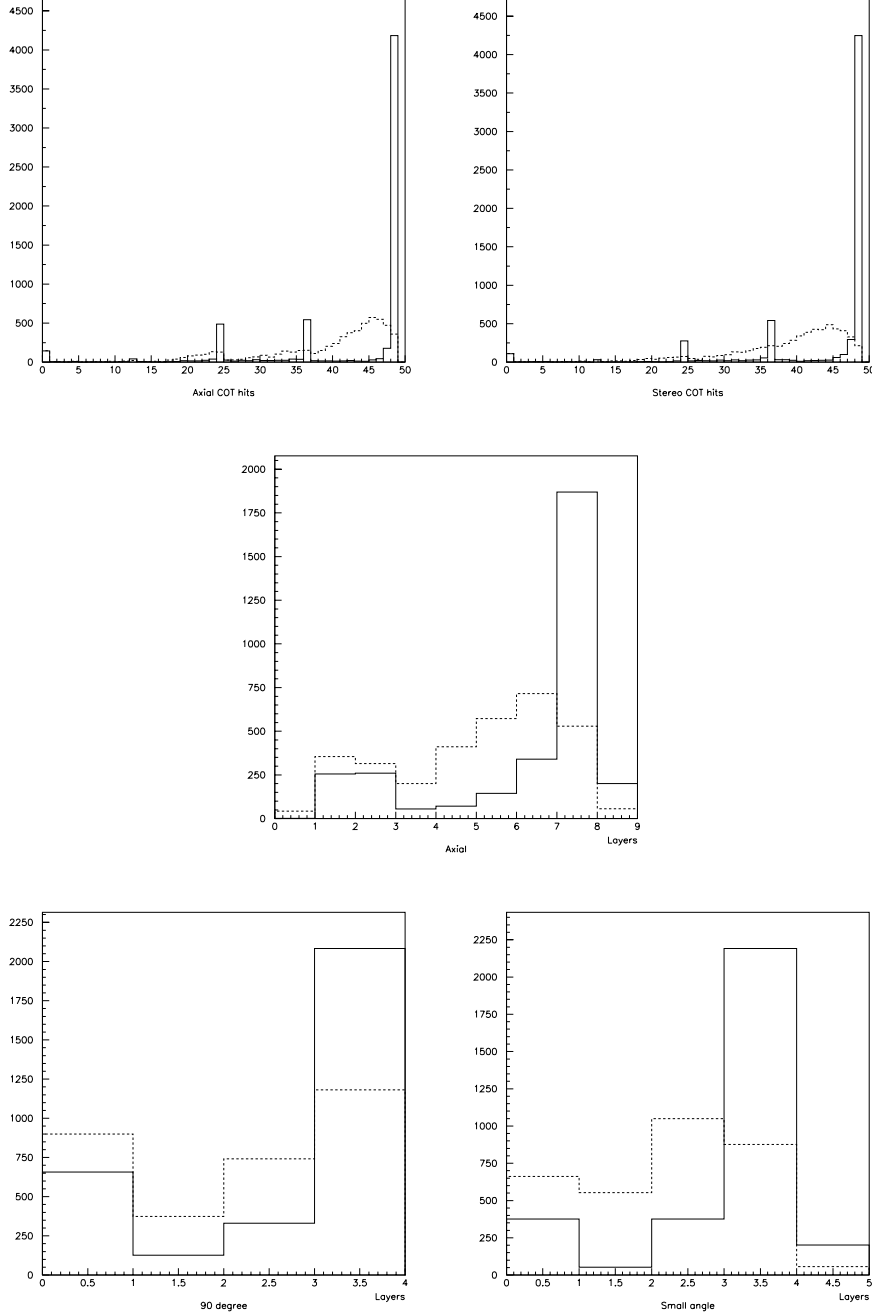
Figure 6.2: Clean (dashed line) and available (solid line) COT hits and silicon layers per track, in $t\bar{t}$ events, divided by type: COT axial (upper left) and stereo (upper right); silicon axial (center), 90° (lower left) and small-angle stereo (lower right).

In order to find a reasonable term of comparison, it is necessary to measure the actual resolution of the tracking system, by histogramming the differences between ideal fit parameters and Monte Carlo parameters, and fitting the result to a gaussian. Since multiple scattering affects low-momentum tracks more than noticeably than stiff ones, the resolution can be profiled as a function of p_t.
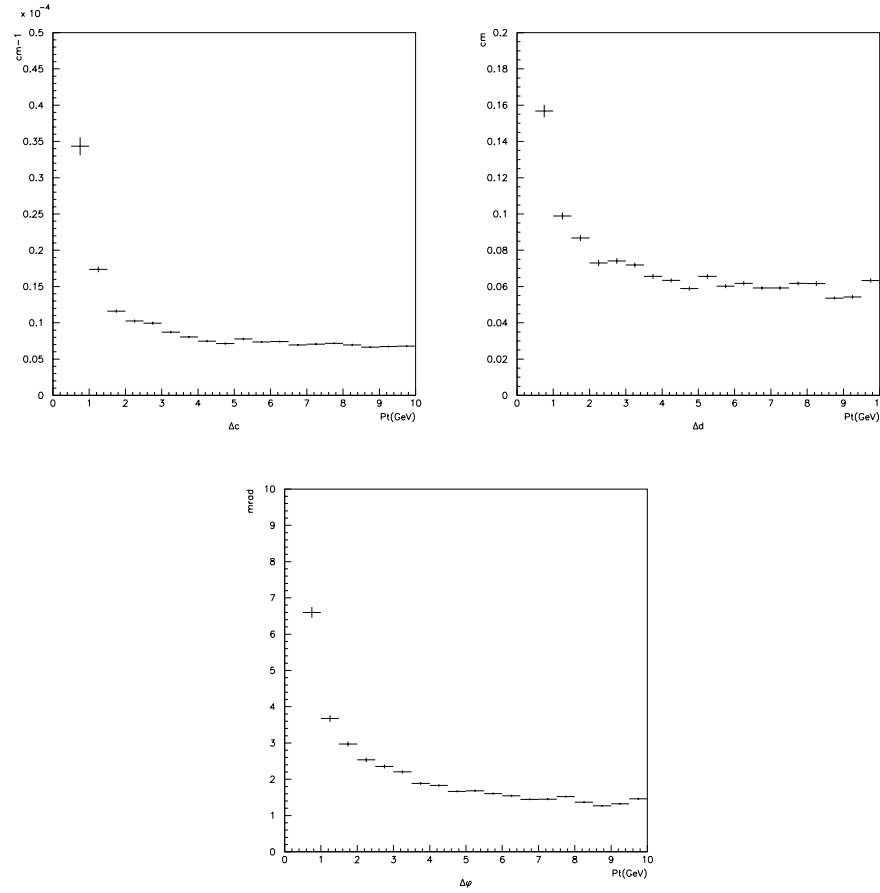
Figure 6.3: Ideal COT resolution after axial reconstruction, as a function of p_t, for single-track events.
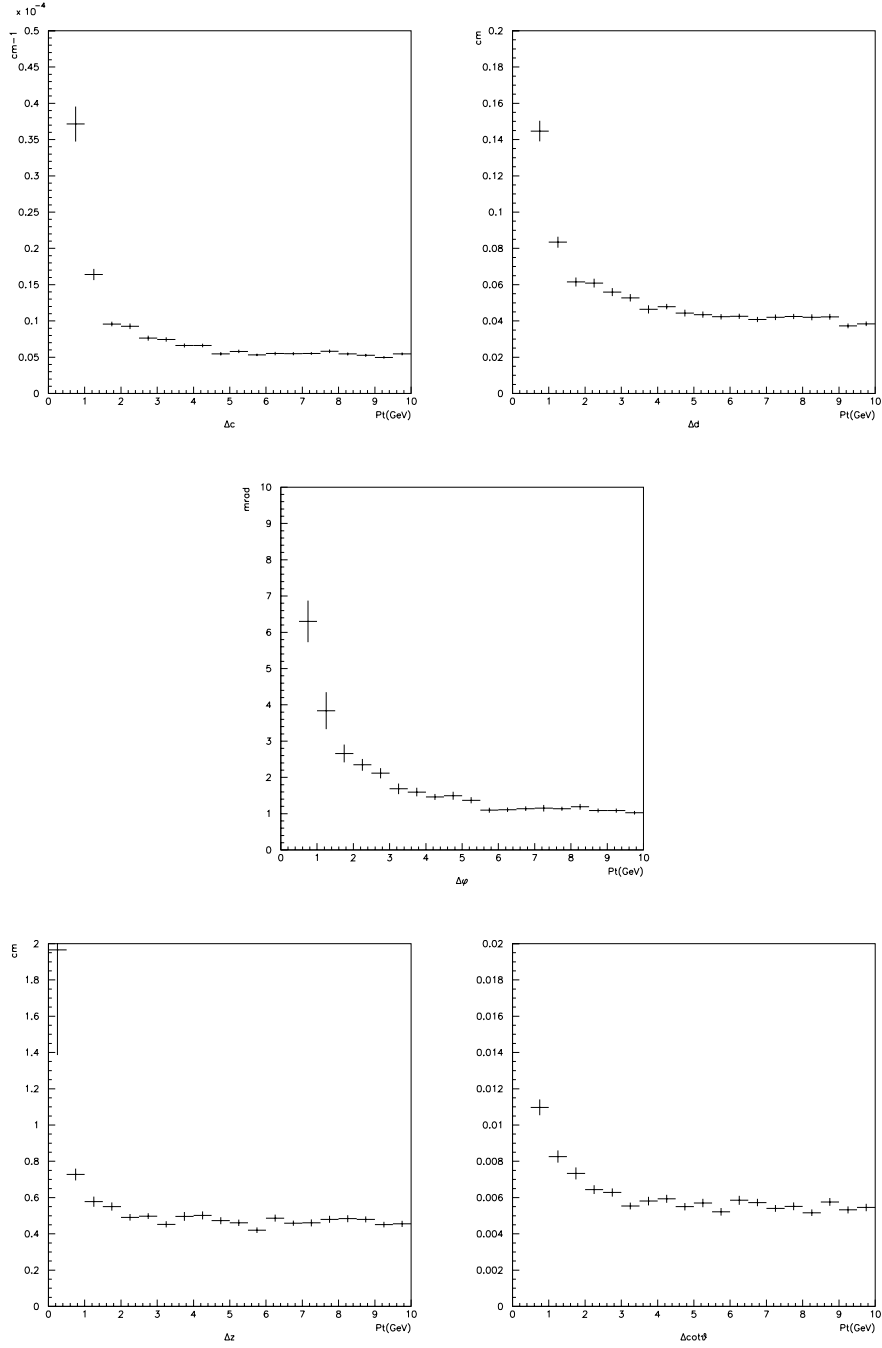
Figure 6.4: Ideal COT resolution after full 3D reconstruction, as a function of p_t, for single-track events.
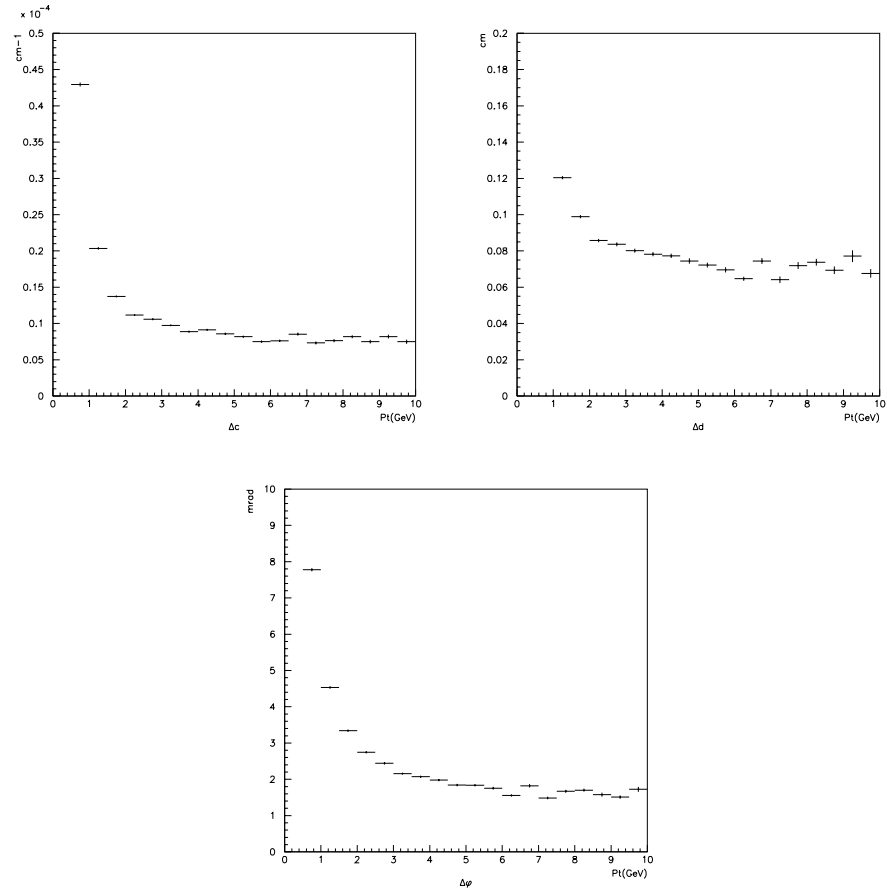
Figure 6.5: Ideal COT resolution after axial reconstruction, as a function of p_t, for $t\bar{t}$ events.
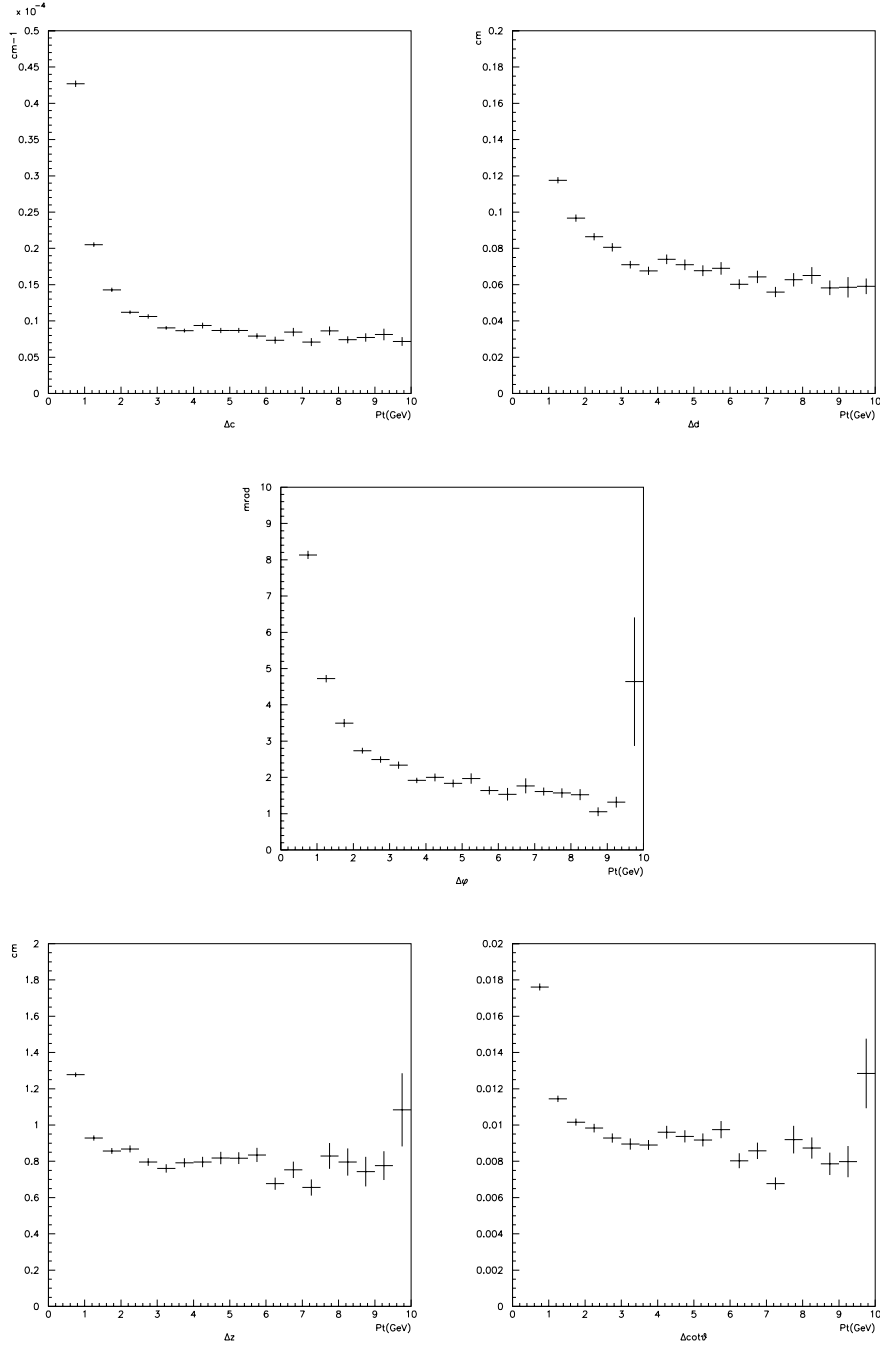
Figure 6.6: Ideal COT resolution after full 3D reconstruction, as a function of p_t, for $t\bar{t}$ events.
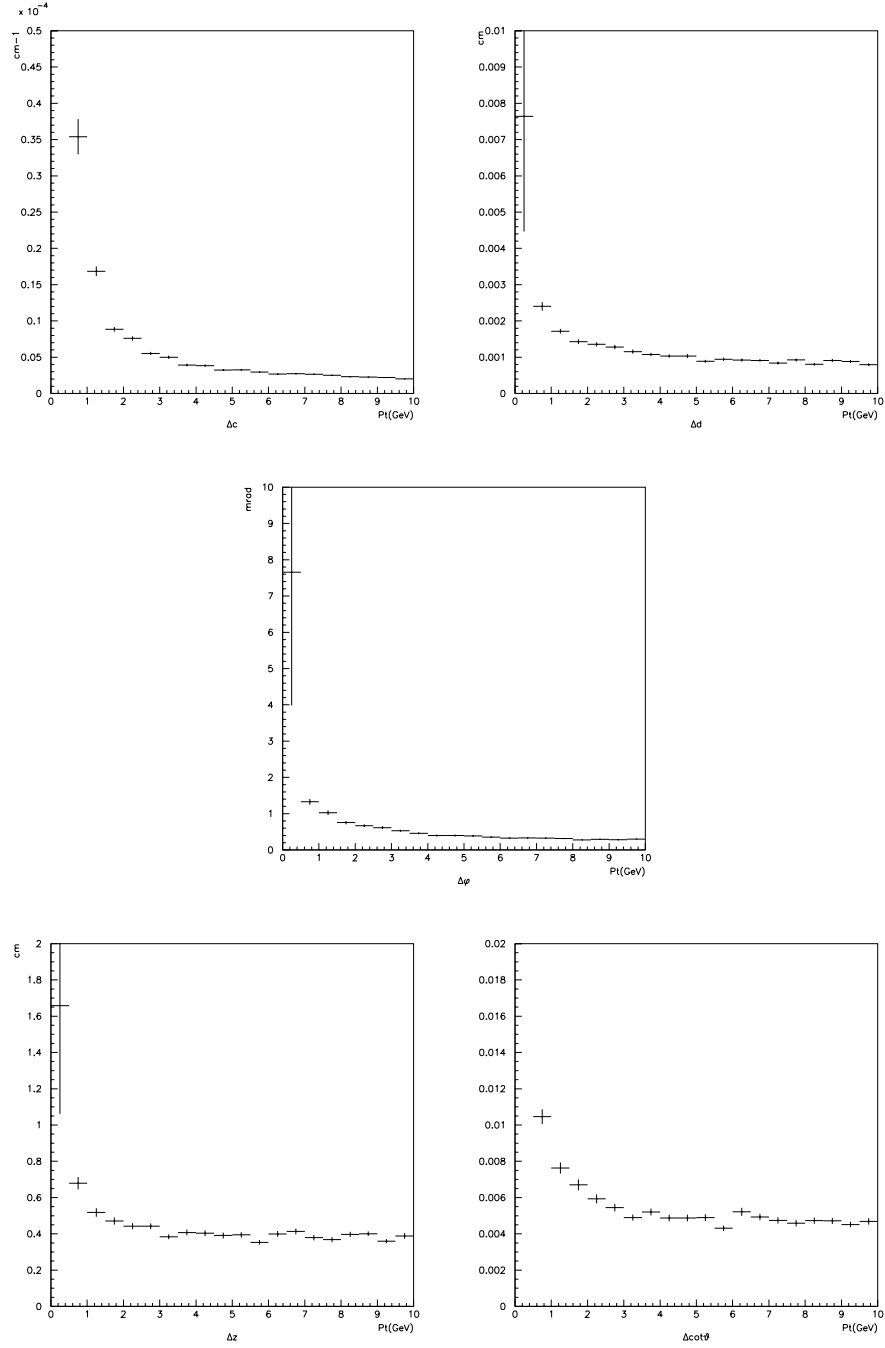
Figure 6.7: Ideal resolution after full COT + axial Si reconstruction, as a function of p_t, for single-track events.
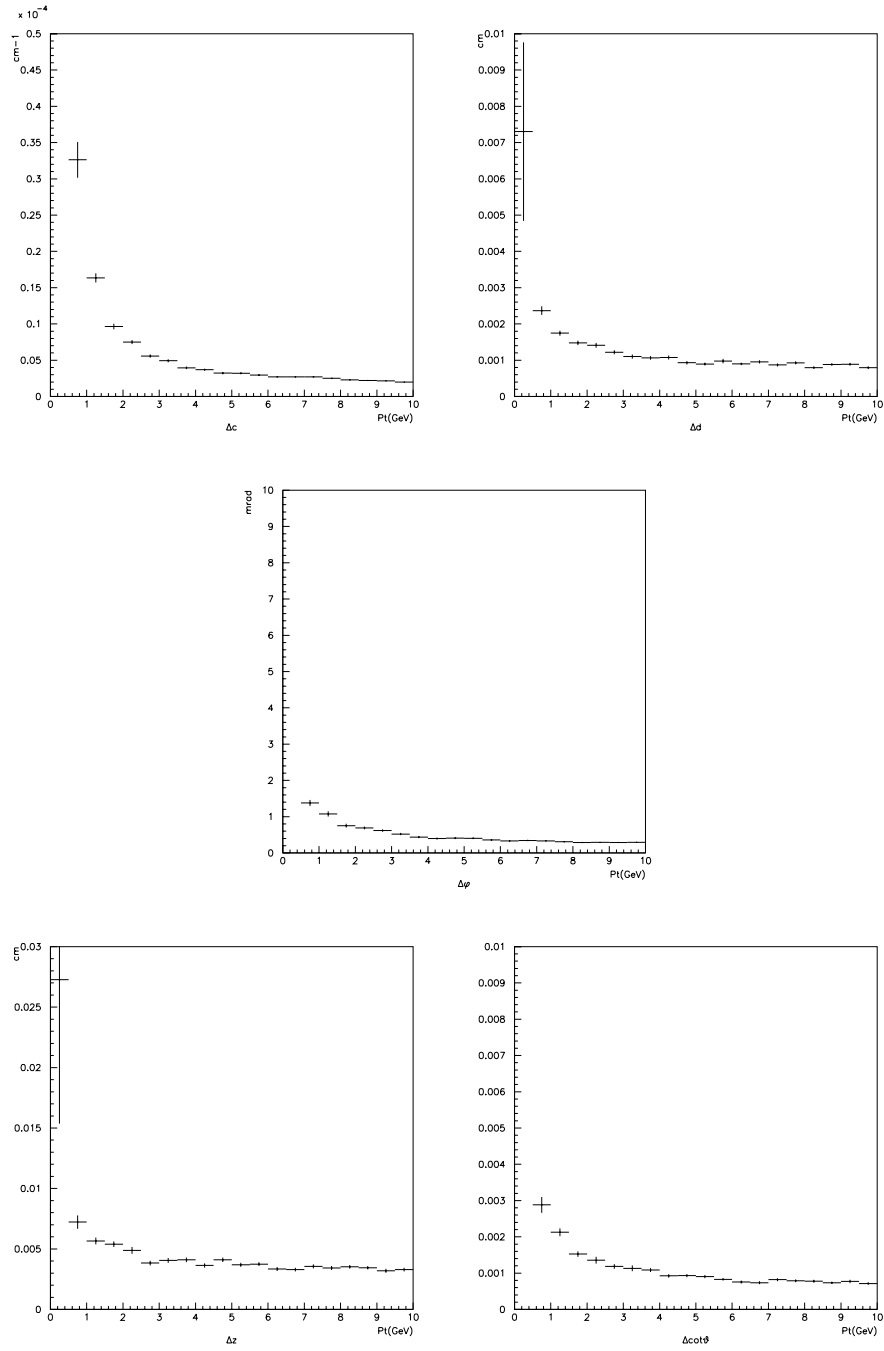
Figure 6.8: Ideal COT+Si resolution after full 3D reconstruction, as a function of p_t, for single-track events.
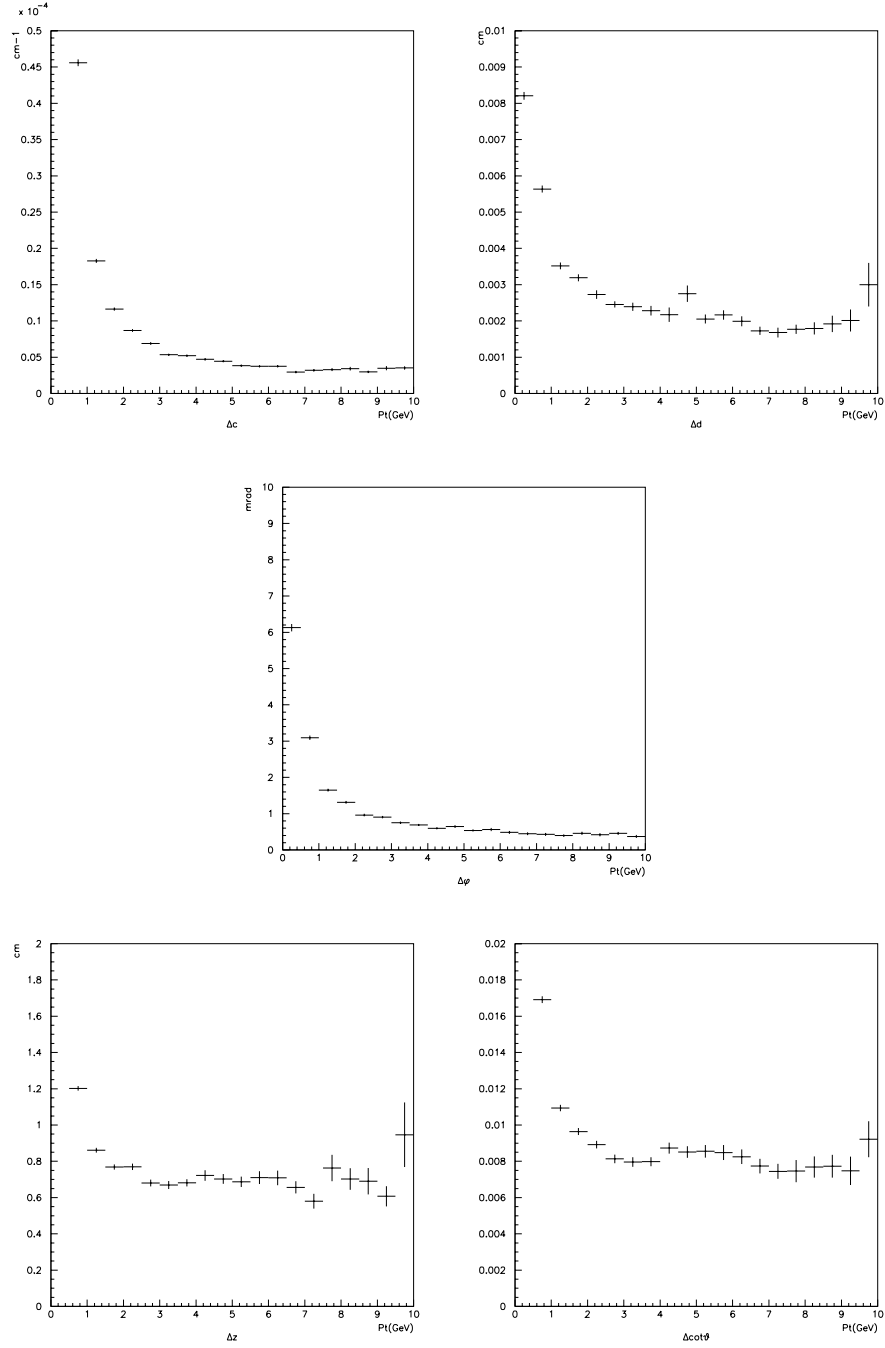
Figure 6.9: Ideal resolution after full COT + axial Si reconstruction, as a function of p_t, for $t\bar{t}$ events.
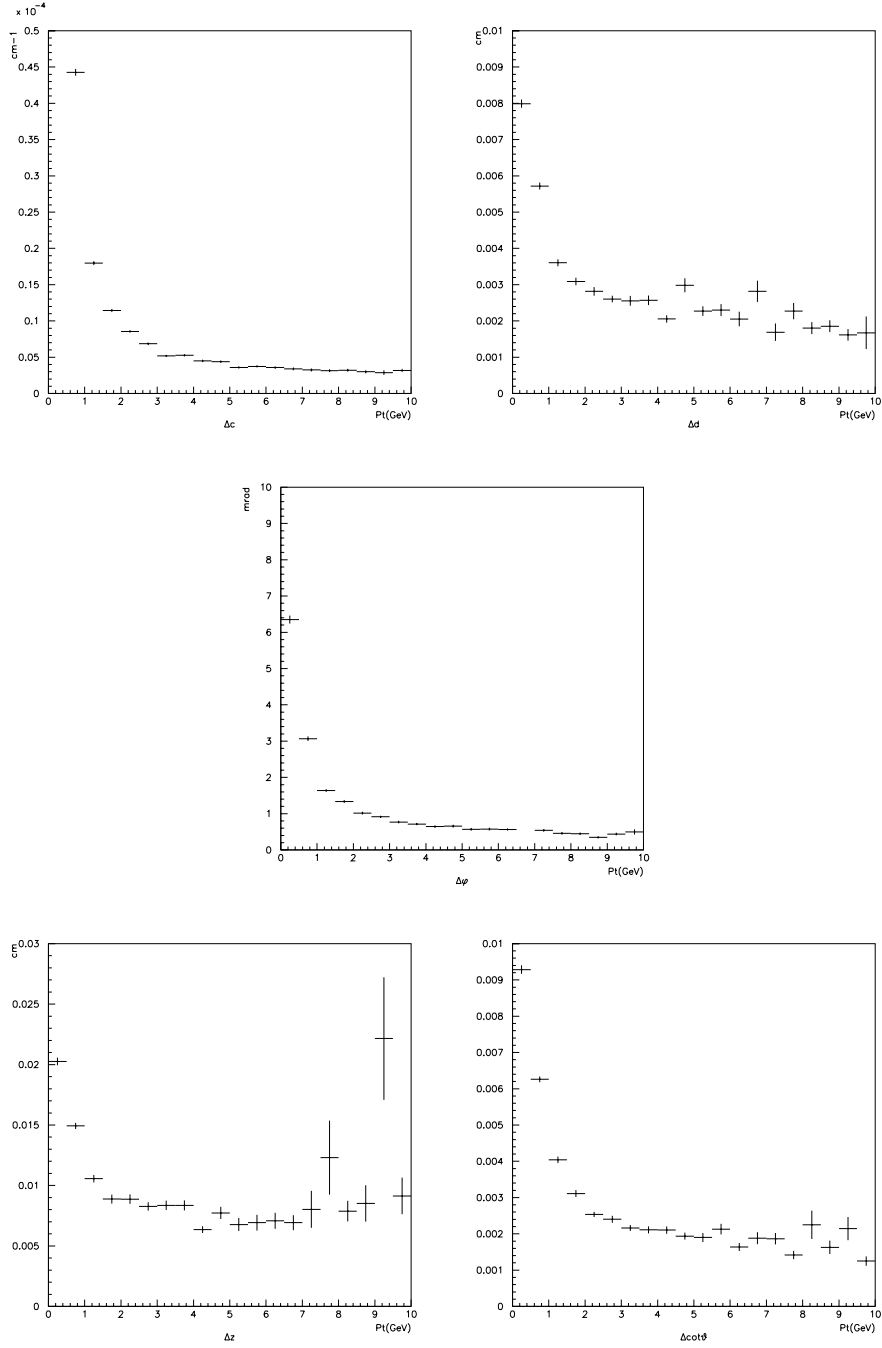
Figure 6.10: Ideal COT+Si resolution after full 3D reconstruction, as a function of p_t, for $t\bar{t}$ events.

Parameter	Sample	p_t	COT 2D	COT 3D	Si 2D	Si 3D
c	μ	1 GeV	$1.7\,10^{-5}$	$1.6\,10^{-5}$	$1.6\,10^{-5}$	$1.6\,10^{-5}$
(cm^{-1})		10 GeV	$0.7\,10^{-5}$	$0.5\,10^{-5}$	$0.2\,10^{-5}$	$0.2\,10^{-5}$
	$t\bar{t}$	1 GeV	$2\,10^{-5}$	$2\,10^{-5}$	$1.8\,10^{-5}$	$1.8\,10^{-5}$
		10 GeV	$0.8\,10^{-5}$	$0.8\,10^{-5}$	$0.3\,10^{-5}$	$0.3\,10^{-5}$
d	μ	1 GeV	1000	800	28	18
(μm)		10 GeV	600	350	8	8
	$t\bar{t}$	1 GeV	1200	1200	35	35
		10 GeV	700	600	18	18
φ_0	μ	1 GeV	3.6	3.6	1.3	1.0
(mrad)		10 GeV	1.0	1.0	0.4	0.4
	$t\bar{t}$	1 GeV	4.6	4.6	1.6	1.6
		10 GeV	1.6	1.2	0.4	0.4
z_0	μ	1 GeV	—	0.75	0.7	0.006
(cm)		10 GeV	—	0.5	0.4	0.003
	$t\bar{t}$	1 GeV	—	0.95	0.7	0.011
		10 GeV	—	0.75	0.5	0.007
$\cot\theta$	μ	1 GeV	—	0.0085	0.0075	0.0022
		10 GeV	—	0.0055	0.0045	0.0008
	$t\bar{t}$	1 GeV	—	0.0115	0.011	0.004
		10 GeV	—	0.0080	0.0075	0.0015

Table 6.2: Summary of ideal resolutions

Results are displayed in figures 6.3 to 6.10, and summarized in table 6.2. The improvement on the various parameters due to each detector part are clearly visible, as well as the dependance on momentum and occupancy.

Even starting from the ideal resolution, the choice of a threshold for fit-based efficiency is somewhat arbitrary. Table 6.3 shows the choices that were actually made; they correspond to about five times the asymptotic resolution for $t\bar{t}$ events, on the relevant parameters of each reconstruction step. These thresholds will be used in Chapters 7 and 8 to evaluate the performance of tracking algorithms.

Parameter	COT 2D	COT 3D	Si 2D	Si 3D
c (cm^{-1})	$4\,10^{-5}$	—	$1.5\,10^{-5}$	—
d (μm)	3500	—	90	—
φ_0 (mrad)	8	—	2	—
z_0 (cm)	—	4	—	0.035
$\cot\theta$	—	0.04	—	0.006

Table 6.3: Fit-based efficiency thresholds

Chapter 7

Tracking in the Central Outer Tracker

This chapter deals with the reconstruction of tracks in the Central Outer Tracker, the drift chamber that was described in Section 3.3.

The reconstruction process begins with the translation of the TDC readout into hit positions; hits are then joined together, to form line segments and tracks. The CDF I track-finding algorithm, called "Segment Linking", is described first; the central part of the chapter introduces a different, "Histogram"-based, solution to the same problem.

The final section is devoted to reporting the performance of the COT algorithms, from two related points of view: time cost and efficiency.

7.1 Drift corrections

The COT hardware provides, as its output, the drift time for each detected hit. In order to perform track reconstruction, this time measurement needs to be converted into a position measurement. First of all, the TDC output t is used to compute the drift distance D; once D is known, the hit position is given by

$$\vec{x} = \vec{x}_w \pm D\hat{v} \tag{7.1}$$

where $\vec{x}_w$ is the wire position, and $\hat{v}$ the drift direction. The $\pm$ ambiguity is due to the fact that the drift sign is unknown: the charge collected by each wire could have drifted clockwise or counterclockwise. This ambiguity is resolved during the process of track finding.

In an ideal detector, the relationship between t and D is trivial: $D = vt$, where v is a constant drift velocity. In the real detector, however, there are

some disuniformities and nonlinearities that need to be corrected. Typical examples that were observed during Run I are:

- The signal propagation delay in the read-out electronics is not identical on all channels. For each channel, t must be offset by an ad-hoc amount, which can be measured during calibration.

- The delay discussed above is not constant during operation. The average change in drift time offset is measured by an online monitor and corrected.

- In the close proximity of sense wires, the drift velocity is not uniform, and is slightly higher than in the rest of the cell. This translates into a correction to very short drift distances.

- Two pulses, starting at the same time but with different amplitude, trigger the readout discriminator at different times and for different durations. To correct this problem, t is adjusted with a term depending on the pulse width.

All of these corrections only depend on the status of individual wires; they are applied as soon as COT data is read into the reconstruction program, so that pattern recognition can benefit from the improved precision. Some other corrections, on the other hand, depend on the parameters of the track that generated the hit. For example:

- Time of flight corrections. A particle propagating with $\eta = 0$ reaches the COT outer layers about 1 ns earlier than a particle with $\eta = 1$.

- "Aspect angle" corrections, depending on the angle between the track and the drift field.

- The luminous region has a RMS spread of about 30 cm along the z axis. This translates in a 2 ns uncertainty on the drift time, that can only be corrected when the z position of the primary vertex is known.

This second group of corrections is applied only after track reconstruction is complete, during the final fitting of track parameters.

7.2   Segment reconstruction

Once the positions of COT hit candidates are known, the tracking code scans each of the eight superlayers looking for line segments. These segments are

used in subsequent steps, either as basic building blocks (in the Segment Linking algorithm) or as reconstruction seeds (in the Histogram Tracking algorithm).

Segment finding begins by looking for triplets of still unused, aligned hits belonging to consecutive layers. Both possible choices of drift sign are taken in consideration for each hit. A list of candidate segments is formed by selecting the cases in which the central hit lies close enough to the midpoint of the external ones, and the overall slope of the segment, α, with respect to the radial direction is not too high.

Candidate segments are arranged in increasing order of $|\alpha|$ (so that high momentum tracks will be given precedence), and fitted to a straight line which is then extrapolated to the other layers of the superlayer. For each layer, there are three possible outcomes:

- A signal leading edge is found within a 20 ns window of the extrapolated line. The hit, with the appropriate drift sign, is added to the candidate segment; the line fit is recalculated.

- There is no leading edge within the window, but the signal from a previous hit extends into the road. This is considered an "obscured" hit.

- No visible or obscured hit is found. A counter of "missed" wires is incremented; if the number of misses ever exceeds the number of good hits, the segment candidate is discarded.

At this point, the residuals on all accepted hits are recalculated; if the worst hit lies outside the acceptance window, it is discarded and the cleanup process starts anew. Whatever hits survive this stage are extremely well aligned; they are fitted once more to a straight line, and the layers on which no hits were accepted are scanned once again.

COT hits must usually belong to a single segment; in case a hit is shared, it is assigned to the segment with the greater number of hits. The only exception to this rule applies to segments which lie entirely within a single cell. In this case, another valid segment ("ghost" segment) is obtained by flipping the drift sign of all the hits.

7.3 Segment fitting

In order to use a reconstructed segment in the subsequent steps of the track-finding process, it is necessary to determine its parameters by fitting the hit positions to a straight line. The reconstruction formulae become simpler by

choosing a coordinate system with the origin in the center of the cell that contains the greater part of the segment. The x axis is oriented radially, as in figure 7.1; the drift direction (even in adjacent cells, with reasonably good approximation) is parallel to y.
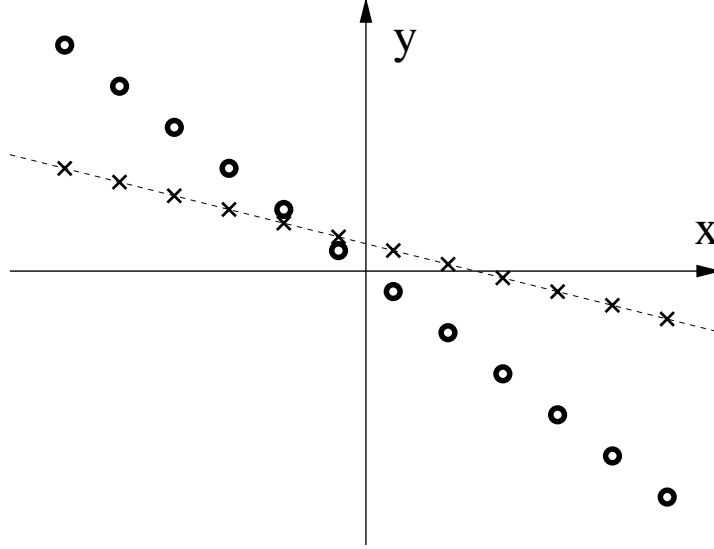
Figure 7.1: Segment fitting within a COT cell. Circles indicate wires, crosses mark the reconstructed hit positions.

The straight-line fit, $y(x) = a_1 + a_2 x$, is performed [64] by calculating

$$
\begin{aligned}
(\Lambda_{11}, \Lambda_{12}, \Lambda_{22}) &= \sum_{i=1}^{12} (1, x_i, x_i^2)/\sigma_i^2 \\
(g_1, g_2) &= \sum_{i=1}^{12} (1, x_i) y_i/\sigma_i^2 \\
D &= \Lambda_{11}\Lambda_{22} - \Lambda_{12}^2
\end{aligned}
\tag{7.2}
$$

The covariance matrix V is the inverse of Λ:

$$
\begin{pmatrix} V_{11} & V_{12} \\ V_{12} & V_{22} \end{pmatrix} = \frac{1}{D} \begin{pmatrix} \Lambda_{22} & -\Lambda_{12} \\ -\Lambda_{12} & \Lambda_{11} \end{pmatrix}
\tag{7.3}
$$

and the fit parameters are

$$
\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = V \begin{pmatrix} g_1 \\ g_2 \end{pmatrix}
\tag{7.4}
$$

The errors σ_i on the y_i hit coordinates are all identical. If it were possible to apply all the hit corrections (discussed in section 7.1) at this stage, σ would be close to 180 μm. Since some corrections can only be performed after track reconstruction is complete, at the level of segment finding σ is larger — about 240 μm (averaged over the entire z extent of the COT).

The covariance matrix does not depend on the y_i values, but only on which terms are summed into Λ — that is, on which layers contribute to the segment. One can therefore determine the segment resolution a priori, as a function of σ, of the segment's radial width Δr (about 6.4 cm), and of the pattern of active wires. In the most common case, where all twelve layers contribute to the segment, Λ_{12} is zero by construction; the covariance matrix is symmetric, and we have

$$\Lambda_{11} \quad = \quad 12/\sigma^2 \tag{7.5}$$

$$\Lambda_{22} \quad = \quad \frac{2}{\sigma^2}\sum_{i=0}^{5}\left(\frac{2i+1}{2}\cdot\frac{\Delta r}{11}\right)^2 \; =$$

$$= \quad \frac{\Delta r^2}{242\sigma^2}\sum_{i=0}^{5}(2i+1)^2 \; =$$

$$= \quad \frac{143\Delta r^2}{121\sigma^2} \tag{7.6}$$

In the best case, then, the fit parameters a_1 and a_2 are totally uncorrelated; their expected errors are respectively

$$\sigma(a_1) = \quad \sigma/\sqrt{12} \quad \simeq 70\,\mu\mathrm{m} \tag{7.7}$$

$$\sigma(a_2) = \quad \tfrac{11}{\sqrt{143}}\tfrac{\sigma}{\Delta r} \quad \simeq 3.4\,\mathrm{mrad}$$

The plots in figure 7.2 show the actual errors measured on each of the four axial superlayers, in a sample of high-p_t single-muon events. Due to the large lever arm of the COT, multiple scattering completely dominates the uncertainty on the center-of-segment position: the segment's distance from the ideal trajectory of the particle is usually much larger than the estimate from eq. 7.7.

7.4   Segment Linking

Once segments are available, the CDF II tracking code tries to assemble them into tracks. The baseline strategy for doing this is called Segment Linking, and is divided in two parts: at first, axial segments are joined in a 2D track; then, stereo segments and individual stereo hits are attached to each axial track.
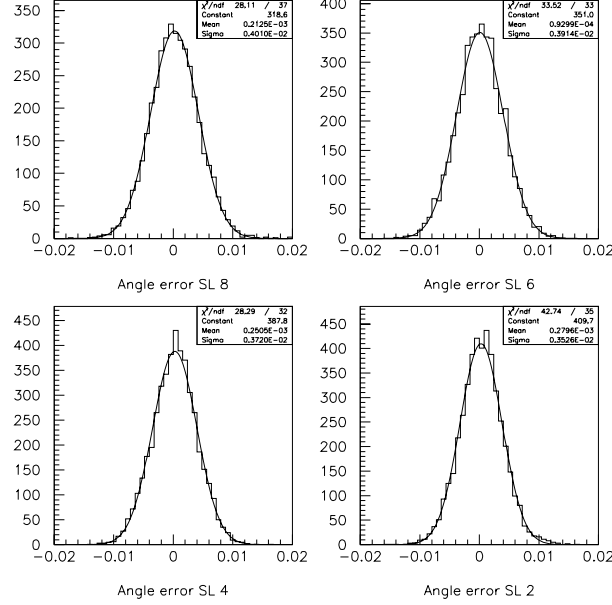
Figure 7.2: Resolution of axial segments in the COT

7.4.1 Axial Segment Linking

Axial Segment Linking is based on a simple geometrical property of the segments that belong to the same track, shown in figure 7.3. If two segments are tangent to the same circumference, they form equal angles α with the line that connects their central points.
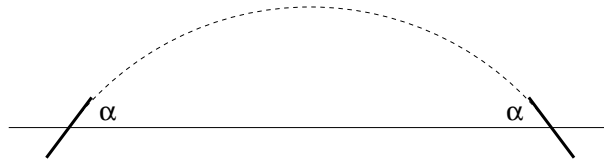
Figure 7.3: Principles of Segment Linking

The algorithm starts taking all the outermost axial superlayer's segments as seeds. Each "outer segment" is extrapolated towards the beam pipe, assuming that the track had a null impact parameter. The intersection of the resulting circle with the second[1] axial superlayer defines the center of a

---

[1] Counting from the outside.

search region, which extends 1 cm on either side. This cut is sufficiently loose to prevent the low-d bias, introduced by the extrapolation, from affecting the final outcome in a relevant way.

All the segments within the selected region of the inner superlayer are then compared to the current outer segment, as shown in figure 7.3. Due to the finite resolution of the detector, the two angles do not match exactly; $S = |\sin \Delta\alpha|$ is used as a figure of merit. If no inner segment matches the outer segment's direction with an S better than 0.025, no linking is performed; otherwise, the lowmost-S inner segment within the search window is linked to the outer segment.

Once all the possible couples of segments on the two outermost superlayers have been linked, the search moves on to the third superlayer. At this stage, two kinds of seeds are used: couples of linked first- and second-superlayer segments, and those first- and second-superlayer segments that were not linked. The latter are handled as in the previous step. As they contain some information on the real curvature of the track, the linked couples are instead extrapolated without having to resort to the $d = 0$ bias.

This procedure is repeated once more to link the innermost axial superlayer's segments to segments and tracks in the outer three superlayers. Finally, the reconstructed tracks which do not contain a segment on the outermost superlayer are extrapolated outwards and matched to any outermost superlayer segments that has not been used yet.

At the end of axial Segment Linking, the found tracks are fitted to circles on the transverse plane. The fit parameters are used in the subsequent 3D pattern recognition phase; should 3D reconstruction fail, the 2D track is still available for event analysis.

7.4.2 Stereo Segment Linking

For each axial track, stereo reconstruction begins by determining which stereo superlayers should be examined: the superlayers that the axial track does not intersect are immediately discarded.

The algorithm then examines the segments on the outermost active stereo superlayer, within 3.5 cells of the axial track. By combining the segment's hits to the axial track parameters, each segment is fitted to a straight line in the rz plane; the segments whose extrapolation intersects the detector axis inside the luminous region are kept.

The following step consists in extrapolating "outer segments" inwards to the next stereo superlayer, and calculating the expected z positions of the intersections. The "inner" segments in the surroundings of this intersection are then examined; each couple of outer and inner segment is fitted to a

single straight line on the rz plane. The segment couple whose χ^2 is lowest (and lower than a predetermined cutoff) is attached to the axial track.

At this point, the current track contains enough 3D information to make a simpler approach possible for the remaining stereo superlayers: the 3D track is extrapolated inwards, segments within a small φ and z window are fitted to the track, and the best-χ^2 segment is attached to the track (once again, provided χ^2 is below threshold).

7.4.3 Stereo Hit Linking

In case stereo Segment Linking fails to find a match on some superlayer, a fallback algorithm is used to attach single stereo hits to the track. This was the default algorithm during Run I: the COT's predecessor only had 6 wires per stereo superlayer, and was unable to perform reliable stereo segment reconstruction.

Before the algorithm starts, the z positions of primary vertices are roughly determined by examining the three 90° layers of the silicon detector. Each aligned triplet of hits is fitted with a straight line and extrapolated to the detector axis; the intersection is stored in a histogram, and the histogram peaks are interpreted as primary vertices.

Like stereo Segment Linking, Hit Linking begins by determining which stereo superlayers should be searched. Beside requiring the axial track to intersect stereo layers, the algorithm needs the track not to form too wide an angle with the radial direction.

At this point several bitmaps are zeroed. Each bitmap contains one row for each stereo layer, and one column for each 1-cm z slice of the detector. Cells surrounding the intersections of the axial track with each stereo super-layer are then examined; for each hit and for both possible drift signs, the z coordinates corresponding to the leading and trailing edge (z_L and z_T) are computed, and used to fill the bins of each bitmap as follows.

- "Wide misses" are the bins that lie more than 5 cm away from any z_L—z_T interval on the current layer.

- "Narrow misses" are defined in the same way, but with a 1 cm threshold.

- "Far hits" require a leading edge within 2 cm of the current bin.

- "Near hits" are like far hits, with a 1 cm threshold.

The algorithm now loops over all the primary vertices, to pick up a z_0 value, and scans the entire $\cot\theta$ range in the bitmap for the "best" road.

The track is required not to have too many misses in either of the first two bitmaps; "far" hits, and then "near" hits, are used to choose between multiple candidates.

Once the best z_0 and $\cot\theta$ combination has been found, actual stereo hits are linked to the track. On each layer, the algorithm selects those hits whose leading edge falls within 2 cm of the predicted position; if more than one candidate is found, the closest one is selected.

7.5 Histogram track reconstruction

The Segment Linking algorithm is based on the standard "divide et impera" problem-solving technique: the global problem ("find all tracks") is divided into a series of smaller problems ("find all segments on each superlayer"), and the partial solutions to these problems are then merged together. As a result, the algorithm is quite fast. However, this speed is achieved at a cost: the global efficiency is extremely sensitive to the small-problem resolution efficiency. With a 1% inefficiency on segment reconstruction, for example, even a perfect axial segment linker will yield 4% of tracks with a wrong or missing axial segment.

High hit-efficiency and purity are particularly important in track finding: missing a segment reduces the fit precision, while attaching a single spurious segment to a track leads to a substantial shift in the fit parameters.

For these reasons, a completely different approach to track finding has been investigated: Histogram Tracking.

7.5.1 Generalities

Histogram Tracking, also called **integrated** tracking, is a general-purpose tracking algorithm. The strategy is seeded by a **telescope**, that is a hint on the position and parameters of a track; it scans several detector **layers**, and determines the set of **hits** that, together with the telescope, form the best track.

In order for the algorithm to be useful, telescopes should describe the track with good accuracy in a subspace $\mathcal{T}$ of the parameter space $\mathcal{P}$; single hits, once associated to the telescope, should improve resolution along a direction $\mathcal{D}$ of the parameter space; $\mathcal{D}$ should be orthogonal, or almost orthogonal, to $\mathcal{T}$. For example, a telescope could consist in a fully reconstructed COT track, providing good curvature and φ_0 measurement, but poor d resolution; once associated to the correct axial hits in the silicon vertex detector, its impact parameter resolution would greatly improve. $\mathcal{D}$ does not need to be

one of the helix parameters; in principle, any function of them can be used. The basic principle of the algorithm is that, by shifting the telescope's extrapolation rigidly along $\mathcal{D}$, at a certain point it should pass "close" to most of the corresponding track's hits.
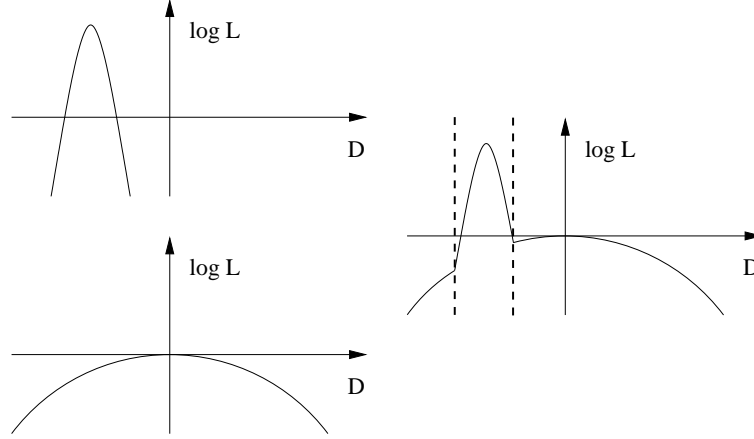
Figure 7.4: Histogram Tracking applied to a single layer and hit. The plots show the likelihood that the error on the current telescope's parameters is $\mathcal{D}$, assuming that the current hit belongs (above) or not (below) to the track. The most likely event is shown to the right.

Once a telescope is known, it is extrapolated into the layers to be scanned, as shown in the left half of figure 7.5, and compared to the set of hits. To speed up the process, only the hits within a reasonably wide search window around the extrapolation are examined.

At this point, the algorithm assumes that all the telescope's parameters are correct, except $\mathcal{D}$; it then determines a likelihood $\mathcal{L}(\mathcal{D})$ for a range of $\mathcal{D}$ values, and selects the point with the maximum likelihood.

If the detector consisted of a single layer, and there were a single hit on that layer, the situation would be particularly simple: either the hit belongs to the track, or it does not. The problem consists in determining which of the two possibilities is the most likely. To obtain the answer, the algorithm plots the likelihood of both cases as a function of $\mathcal{D}$, as in figure 7.4, into a histogram; in each bin, the highest of the two contributions is selected. The global maximum of the histogram corresponds to the most likely value of $\mathcal{D}$; once this is known, one can determine whether the hit should or should not be associated to the telescope.

The situation does not become much more complicated if the layer contains multiple hits: once again, the algorithm has to plot the likelihood of

each event (does the track pass through hit 1, through hit 2... or through no hit at all?) as a function of D, select the most likely event in each bin, find the global maximum, and choose the eventuality corresponding to the most likely event in the best D-bin.
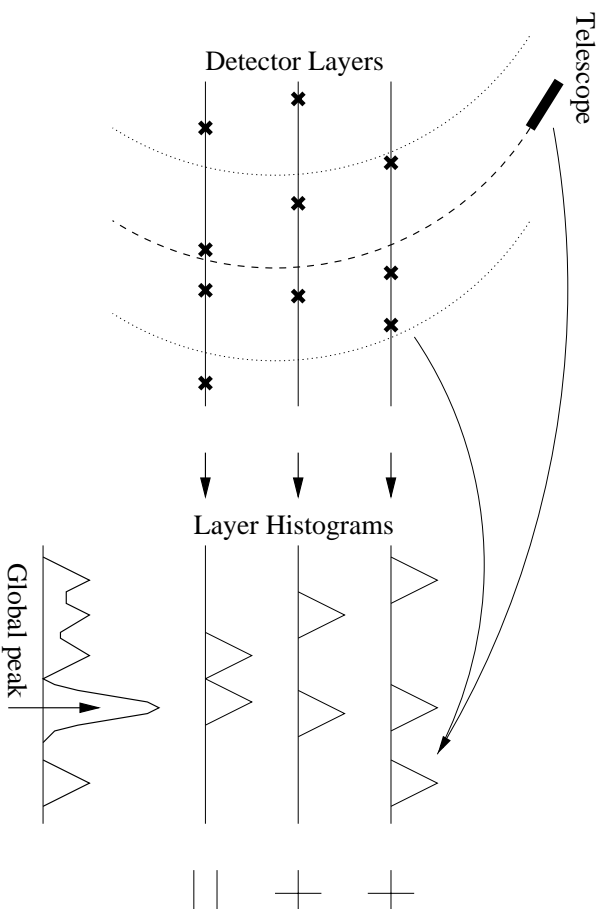
Figure 7.5: Schematics of Histogram Tracking. The thick line indicates a telescope, the dashed line its extrapolation, the dotted lines the corresponding search window, and crosses are detector hits. Each hit is converted to a contribution to the histograms in the right half of the picture; the histograms are then summed, and the position of the global maximum is used to select hits belonging to the track.

The situation becomes more interesting when the detector consists of multiple layers. In this case, the global likelihood is no longer associated to the best single hit, but to the best **combination** of hits on the various layers. To calculate $\mathcal{L}$, the algorithm determines single-layer likelihoods as before; the layer contributions are then multiplied (or rather, for the sake of speed, their logarithms are summed) bin by bin, as in the right half of figure 7.5. Then, as usual, the algorithm seeks the global maximum of $\mathcal{L}$; once the best value of D is known, the most likely combination of hits can be attached to the telescope.

Contrarily to other strategies, like Segment Linking, Histogram Tracking operates at the level of single hits. This increases the time spent during reconstruction, as more comparisons are required. On the other hand, it also

increases efficiency: the algorithm can succeed even in those cases when the intermediate segment-finding step would fail.

Histogram Tracking builds its histograms independently on each layer; it is therefore well suited to parallel processing. The partial results are then summed[2], giving equal importance to all the candidate hits. As a consequence, the algorithm is not particularly vulnerable to the presence of isolated, spurious hits: only a coincidence between several layers can seriously damage the end result.

A final characteristic of integrated tracking is its behaviour at high occupancy. The number of telescopes is usually proportional to the occupancy O of the event; the number of hits within the search window is also proportional to O, making the total search time proportional to O^2. Moreover, the elementary step (determining the likelihood plot for each hit/telescope combination) does not usually require slow operations. Other algorithms, such as the "Outside-In Tracking" described in chapter 8, perform nested loops on each layer's hits, and have a very slow inner loop, entailing operations such as helix fits; the search time grows with a high power of occupancy. As a consequence, Histogram Tracking — while being "slow" on simple events — can be quite fast on complicated ones.

Implementation details

In order to improve performance and execution speed, the algorithm actually used at CDF II differs slightly from the previous description.

First of all, the global histogram maximum is required to exceed a certain threshold; if the peak is too low, the algorithm assumes there was no track to be found. This is done to discard "fake" telescopes that do not actually correspond to a track. Almost every telescope matches a few hits for certain values of $\mathcal{D}$; this cut makes sure there is a minimal number of excellent matches, or a higher amount of approximate matches.

While filling the single layer histograms, the no-hit contribution to the likelihood is not computed as a function of $\mathcal{D}$, but it is assumed to be constant within the $\mathcal{D}$ range of the histogram; it is then possible to rescale all plots, and set this constant equal to zero. Due to the peak height cut, this approximation does not influence the final outcome at all; on the other hand, filling the histogram with a constant is much faster than performing a calculation for each bin.

Single hit contributions to the layer histogram are also calculated with an approximation. Instead of determining $\mathcal{L}$ in each bin as a function of

---

[2]The sum of contributions from different parts of the detector is the reason why this algorithm is also called "integrated" tracking.

the hit and telescope parameters, the algorithm finds the value of $\mathcal{D}$ corresponding to the maximum likelihood for the current hit; it then expands $\log \mathcal{L}_{\text{hit}}$ in a Taylor series around its maximum. Since the no-hit likelihood is approximated with a constant, it is also easy to find the range of histogram bins (which is usually quite small) where the hit contribution prevails. As a consequence, the histograms can be filled very quickly.

Finally, once the hits contributing to the histogram maximum have been found, the algorithm performs some cleanliness cuts. For example, if on a certain layer there are two hits equally close to the global peak, the histogram has to decide whether to accept both, either, or none; the details of this decision depend on the specific situation.

7.5.2 Histogram Tracking in the COT

Histogram Tracking has been applied to axial track reconstruction in the COT. Two fundamental choices of telescope are possible in principle: either a single segment, or a couple of pre-linked segments. However, once two segments have been cleanly linked, the most difficult part of Segment Linking is complete. Thus, switching to Histogram Tracking at this point would not bring major advantages.

Seeding the algorithm with a single axial segment, and scanning the other three axial superlayers for hits, is a more promising approach. However, single segments contain very little information: as was shown in section 7.3, their direction is known with poor resolution, and only their center's position can be measured accurately. In the crowded environment expected for Run II, this is a major obstacle.

The solution to this problem consists in taking advantage of the typical geometry of events. A large fraction of the physically relevant tracks originate in the primary interaction(s); several other tracks start in secondary vertices placed inside the beam pipe. It is therefore possible to perform a first round of Histogram track reconstruction, seeded with single axial segments, with the added constraint that the track's impact parameter should be small. Once the majority of tracks has been found, and their hits have been removed from the candidate list, the occupancy is small enough for the beam constraint to be released.

In both cases, once the peak in the global histogram is located, selecting hit candidates to associate to the track is a simple process. Layers are scanned one at a time. Within each layer, if only one hit candidate contributed to the global maximum, that hit is associated to the track. If multiple hits contributed, the one with the highest contribution is chosen.

This simple approach is justified by the fact that, with its 36 layers (plus

12 in the telescope), COT Histogram Tracking is not very sensitive to isolated mistakes: selecting a wrong hit on a single layer does not affect the final fit parameters much, and therefore does not degrade performance noticeably. The situation is quite different in the silicon detector, as will be explained in section 8.6.

7.5.3   Constrained telescope

A "constrained" COT telescope consists of an axial segment, reconstructed as explained in section 7.2, and "forced" to belong to a low-d track.
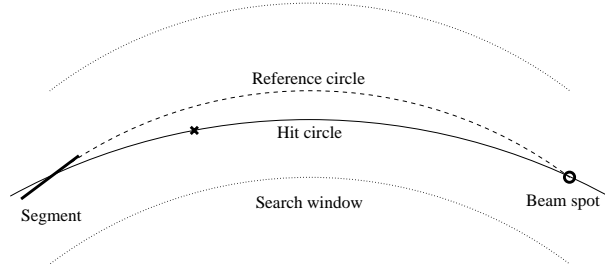
Figure 7.6: Beam-constrained COT telescope: single hit

The seed segment's position and direction are used to determine a "reference circle", passing through the beam spot; on each axial superlayer, hits within 2 cm from the reference circle are examined. For either drift sign, a circle is then drawn through the beam spot, the hit, and the midpoint of the seed segment, as in figure 7.6. As was stated above, the segment's direction has an often large error. In this second step, therefore, the direction is neglected: the search window is narrow enough, and the COT lever arm long enough, to make its further use unnecessary.
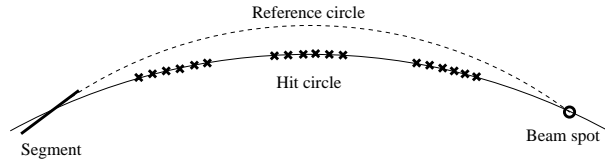
Figure 7.7: Beam-constrained COT telescope: multiple hits

As shown in figure 7.7, hits belonging to the same low-d track lie on the same "hit circle", within a margin of error that depends on the single hit resolution and on the effects of multiple scattering. Hits belonging to other

tracks, or candidate hits with the wrong drift sign, usually lie at a greater distance from this circumference. Therefore, one can choose, as abscissa for the integrated tracking histograms, the parameter difference between the reference circle and the hit circle. In the actual implementation, in order to make the plots easier to visualize, this difference is plotted as the azimuthal arc length between the two circles, at a fixed value of r (the "reference radius" r_R).

The single-hit likelihood is approximated with a gaussian; it is maximal at the value of $\mathcal{D}$ for which the hit circle passes exactly through the hit candidate. Since the segment direction is neglected, and since there is always a unique hit passing through any three points, all the single-hit peaks are given the same height.

To determine the best value of $\mathcal{D}$, let r_H and φ_H be the polar coordinates of the hit, and r_S and φ_S those of the segment's center. In the approximation where cr_S and cr_H are small, which holds true for physically relevant tracks, the hit circle's curvature is given by the ratio

$$c \simeq \frac{\varphi_S - \varphi_H}{r_S - r_H} \tag{7.8}$$

and its direction at the origin is

$$\varphi_0 = \varphi_S - \arcsin(cr_S) \simeq \varphi_S - cr_S \tag{7.9}$$

The hit circle's φ at the reference radius r_R is therefore

$$\varphi_R = \varphi_0 + \arcsin(cr_R) \simeq \varphi_S + (\varphi_H - \varphi_S)\frac{r_S - r_R}{r_S - r_H} \tag{7.10}$$

and the most likely value for $\mathcal{D}$ is

$$\mathcal{D}_{best} = r_R \left(\varphi_S + (\varphi_H - \varphi_S)\frac{r_S - r_R}{r_S - r_H} - \varphi_{ref} \right) \tag{7.11}$$

where φ_{ref} is the φ coordinate of the reference circle at radial coordinate r_R; it gives the same contribution to all hits, and its only purpose is to center the histograms around zero.

If σ_H is the single hit position uncertainty, the corresponding error on φ_H is σ_H/r_H; neglecting the error on the center of segment position φ_S, the width of the likelihood gaussian is then

$$\sigma_D = \frac{r_R}{r_H}\frac{r_S - r_R}{r_S - r_H}\sigma_H \tag{7.12}$$
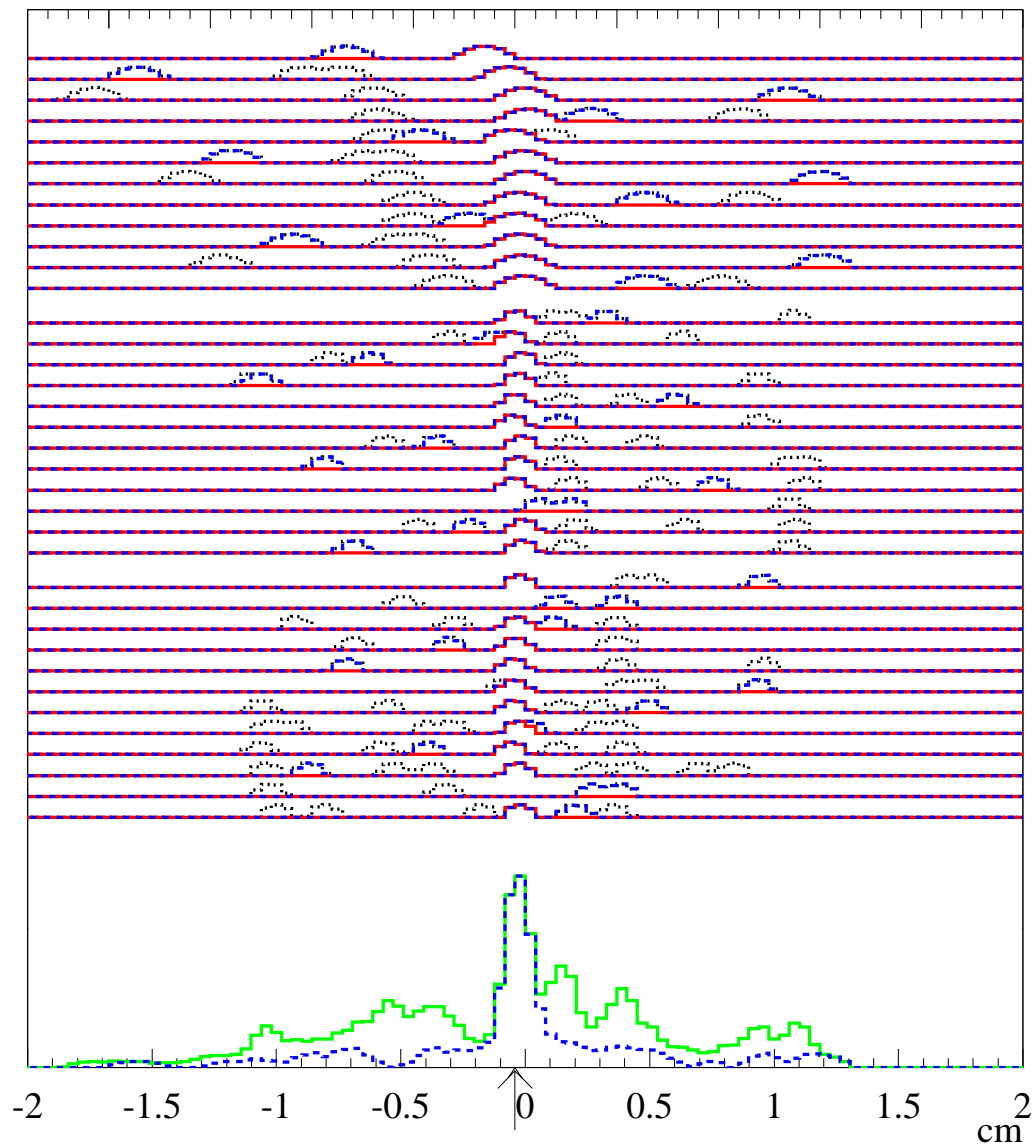
Figure 7.8: Example of Histogram Tracking in the COT. In the upper part of the figure, black dotted lines correspond to hits not belonging to the track; dashed blue lines to hits belonging to the track, but not accepted (most likely because of the wrong drift sign); dashed blue-red lines correspond to the hits correctly associated to the track. In the lower part, the green line is the sum histogram; the dashed blue line is the contribution from correct hits.

7.5.4 Unconstrained telescope

"Unconstrained" COT telescopes only use information from a single axial segment; they do not depend on the beam position, except for the fact that they scan a region of the COT within a (wide) search road centered on a $d = 0$ reference circle, as shown in figure 7.9. Instead, they rely on the segment's position **and** direction to draw hit circles: such circles are all tangent to the segment in its midpoint, and pass exactly each through the corresponding hit.
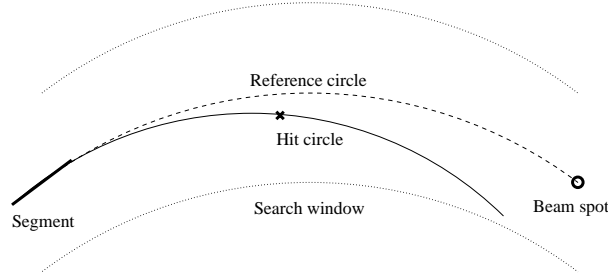
Figure 7.9: Unconstrained COT telescope: single hit
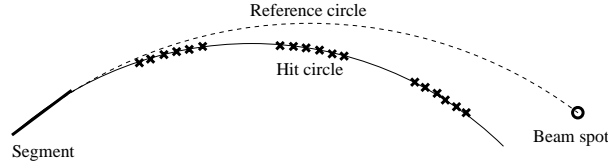
Figure 7.10: Unconstrained COT telescope: multiple hits

As before, hits belonging to the same track lie on compatible hit circles (figure 7.10). The histogram abscissa can be chosen to be the curvature difference Δc between the hit and reference circles; as an alternative, Δc can be multiplied by the square of a constant length ρ_R to obtain a physical, curvilinear distance.

In a polar coordinate system (ρ, ψ) centered on the seed segment, the equation of a circumference tangent to the segment is

$$\psi = \psi_S + \arcsin(c\rho) \tag{7.13}$$

where ψ_S is the direction of the segment. If, in this coordinate system, a hit is found in (ρ_H, ψ_H), then the corresponding curvature is

$$c = \frac{\sin(\psi_H - \psi_S)}{\rho_H} \tag{7.14}$$

and the best value of $\mathcal{D}$ is

$$\mathcal{D}_{best} = \rho_R^2(c - c_{ref}) \tag{7.15}$$

Since the hit circle is unique, there is no parameter of merit on which the peak likelihood could depend on. Therefore, all peaks are given the same height, as in the constrained telescope.
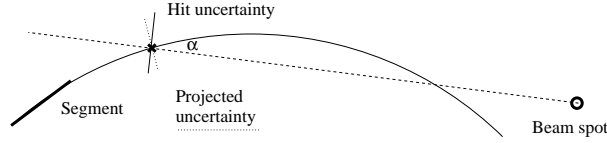
Figure 7.11: Unconstrained COT telescope: error due to hits

If the azimuthal single-hit uncertainty is σ_H, and the hit circle makes an angle α with the radial direction (as in figure 7.11), the error on c is simply $\cos\alpha\,\sigma_H/\rho_H^2$, which can be rounded up to σ_H/ρ_H^2 since α is usually "small" enough for its cosine to be close to 1. The single-hit contribution to the $\mathcal{D}$ uncertainty is then

$$\sigma(\mathcal{D}_H) = \sigma_H\frac{\rho_R^2}{\rho_H^2} \tag{7.16}$$

The best $\mathcal{D}$ values are also systematically shifted due to the error on the seed segment, which is dominated by the direction uncertainty $\sigma(\psi_S)$. Unfortunately, the amount of this shift depends strongly on the hit's layer, being inversely proportional to the segment-hit distance:

$$\delta\mathcal{D} = \sigma(\psi_S)\frac{\rho_R^2}{\rho_H} \tag{7.17}$$

or, by using equation 7.7 for a 12-hit seed segment,

$$\delta\mathcal{D} = \sigma_H\frac{11}{\sqrt{143}}\frac{\rho_R^2}{\Delta r\,\rho_H} \tag{7.18}$$

The segment width Δr can be much smaller than the segment-hit distance ρ_H; therefore, this shift can be much larger than $\sigma(\mathcal{D}_H)$ (eq. 7.16).

For the Histogram Tracking algorithm to succeed, there must be a histogram bin to which all the hits of the track give a non-negligible contribution. Due to the inhomogeneous shift discussed above, to achieve this it is necessary to widen the hits "by hand", according to their ρ_H. As a consequence, unconstrained telescopes are not very effective in a high-occupancy

environment, where multiple hits could be merged together by the widening. This problem can be solved by a preliminar detector clean-up: once the low-d tracks have been found by other strategies, and their hits have been marked as "already used", unconstrained telescopes have a reasonable chance of joining the remaining hits into high-d tracks.

7.6 Performance

The tracking algorithms discussed above were tested on single-track and $t\bar{t}$ samples, in order to estimate their hit-based and fit-based performance, and their CPU usage.

CPU usage was measured on `fcdfsgi2.fnal.gov`, the 64-processor SGI O2000 IRIX mainframe that is used for CDF II offline analysis and development, on the $t\bar{t}$ sample, using the non-optimized development code. The CDF II level 3 trigger is implemented as a scalable farm of Linux PC. It will use aggressively optimized code, and will resort to **regional tracking** — only the most interesting parts of the detector, as determined by the previous levels of trigger, will be examined. The actual performance at trigger level will therefore vary with time, as faster computers are added to the farm and thresholds are adjusted. Nevertheless, the results shown in table 7.1 provide insight on the relative time cost of strategies, and make evident the impact of histogram binning on execution speed.

Operation	Avg time per $t\bar{t}$ event (ms)
Loading hits from TDC	99 ± 2
Axial segment reconstruction	81 ± 3
Axial segment linking	64 ± 2
Axial histogram tracking	
— 50 bins	495 ± 17
— 100 bins (default)	540 ± 20
— 200 bins	640 ± 20
Stereo segment reconstruction	103 ± 3
Stereo segment linking	120 ± 6
Stereo hit linking	75 ± 4
3D track fitting	350 ± 20

Table 7.1: COT reconstruction: CPU timing

From this table, Histogram Tracking appears to be much slower than axial Segment Linking: this seems to rule it out as a possible part of the

level 3 trigger. On the other hand, reconstruction speed can benefit from the use of regional tracking. A full-COT cost of 540 ms might be reasonable, provided the gain in efficiency is relevant.

Figures in the following pages show the tracking efficiency of both strategies on single-track and $t\bar{t}$ samples; table 7.2 summarizes the global efficiency. As expected, the Segment Linking algorithm performance suffers considerably from increased occupancy; Histogram Tracking, on the other hand, performs remarkably well even on $t\bar{t}$ events.

Global efficiency (%)	single tracks ($d = 0$)	$t\bar{t}$ events
Axial:		
— Segment Linking	98 ± 1	87.1 ± 0.6
— Histogram Tracking	99 ± 1	97.8 ± 0.6
Stereo linking:		
— Minimal requirements	99 ± 1	98 ± 1
— Max 1 missed superlayer	98 ± 1	95 ± 1

Table 7.2: Global tracking efficiency in the COT

Efficiency was defined as the ratio between the number of "found" and "findable" tracks; this ratio can be calculated globally, as reported in table 7.2, or as a function of track parameters, as in figures 7.15 and following. "Findable" tracks satisfy the following requisites:

- Originate within the beam pipe[3] and end outside the COT;

- Traverse all of the COT's layers ($|z| < 150$ cm for $r < 140$ cm);

- A sufficient number of clean hits (≥ 20);

- Transverse momentum above 0.5 GeV;

- For stereo efficiency, the corresponding axial track must have been "found".

"Found" axial tracks are the findable tracks that satisfy the fit-based efficiency cuts chosen in table 6.3.

As can be seen in figure 7.15, the constrained telescope can be used with optimal performance on isolated tracks with a real impact parameter up to about 7 mm; this comprises most of the beam pipe's volume, and a vast

---

[3]This cut was relaxed for the sample of high-d single tracks.

majority of the "physically most interesting" tracks. With high occupancy, as in figure 7.16, constrained Histogram Tracking still performs better than Segment Linking up to $d \simeq 5$ mm.

Unconstrained telescopes, although implemented within the Run II tracking framework, are not yet ready to be used. It is necessary to perform further studies on constrained telescopes and Segment Linking, in order to determine how thoroughly the events can be "cleaned"; once this data is known, it will be possible to optimize the unconstrained telescope χ^2 cut. Preliminary studies on single track events show that the efficiency of unconstrained telescopes is substantially flat on an impact parameter range of several centimeters; the overall efficiency exceeds 90% (before any optimization).

The efficiency requirements for associating stereo hits to an axial track are less stringent: while measurement of the track's curvature relies mostly on COT axial data, resolution on the other parameters improves dramatically once the track is extrapolated to the silicon. For silicon reconstruction to be possible, however, the COT track parameters must be sufficiently precise: it is necessary to determine which silicon wafers have been crossed by the track. This requirement can be expressed in terms of hit efficiency in two ways. One consists in asking the track to contain at least two good segments (20+ good hits: the minimum amount necessary to have a reasonable estimate of the track slope and intercept in the rz plane) and few (at most two) wrong hits. Alternatively, the track might be required to have "missed" at most one of its stereo superlayers: there must be at most two wrong and ten missed stereo hits.
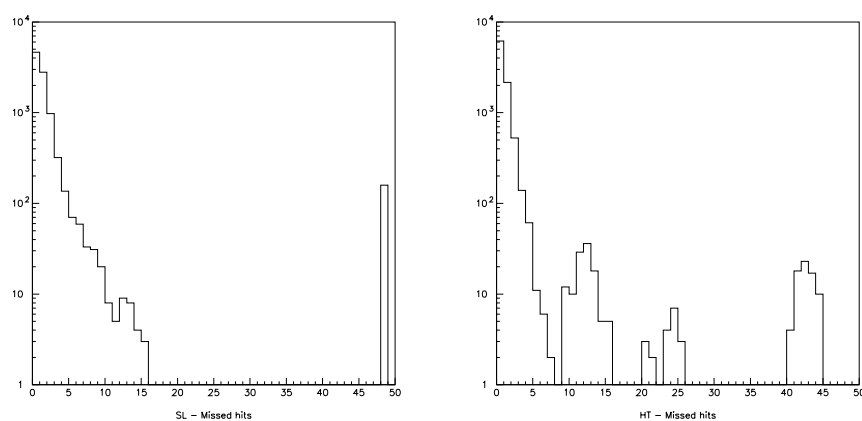
Figure 7.12: Missed hits in COT axial reconstruction on single track events, as performed by Segment Linking (left) and Histogram Tracking (right). Tracks are required to have at least 40 clean hits.
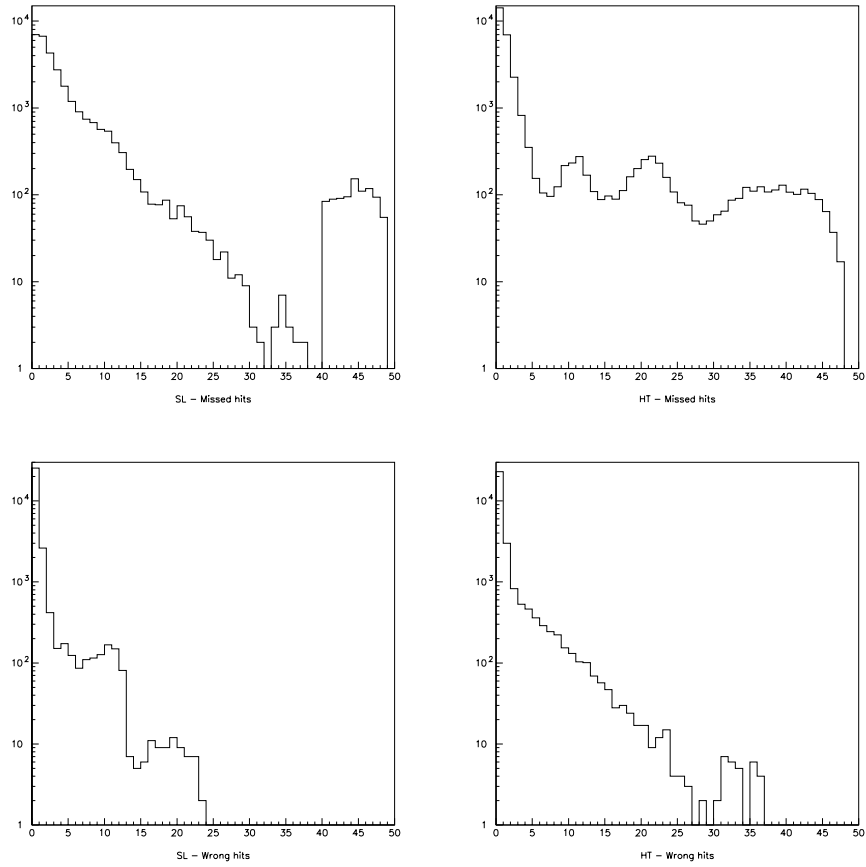
Figure 7.13: Missed (above) and wrong (below) axial hits in COT reconstruction on $t\bar{t}$ events, as performed by Segment Linking (left) and Histogram Tracking (right). Tracks are required to have at least 40 clean hits.
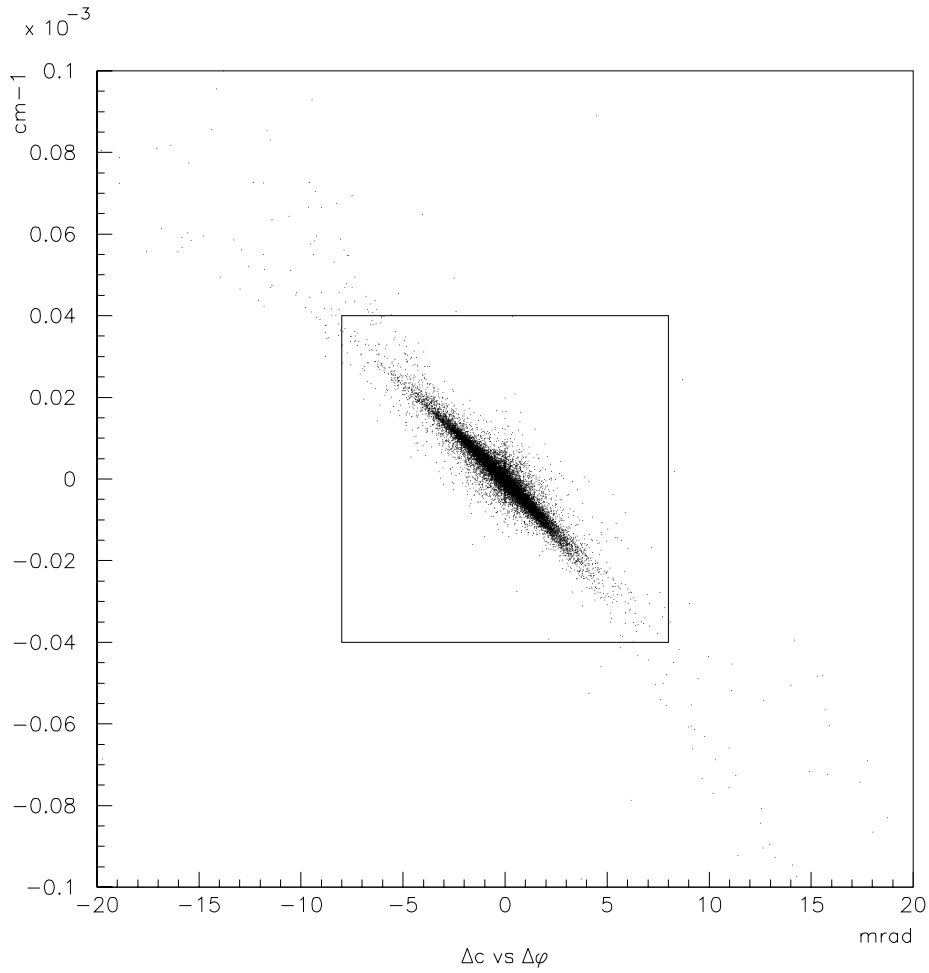
Figure 7.14: COT fit-based efficiency is defined as the fraction of fiducial tracks (i.e. tracks within the COT fiducial volume) for which the difference between the real fit and the ideal fit parameters does not exceed a given threshold, represented by the inner box. Points outside the box correspond to badly reconstructed tracks.
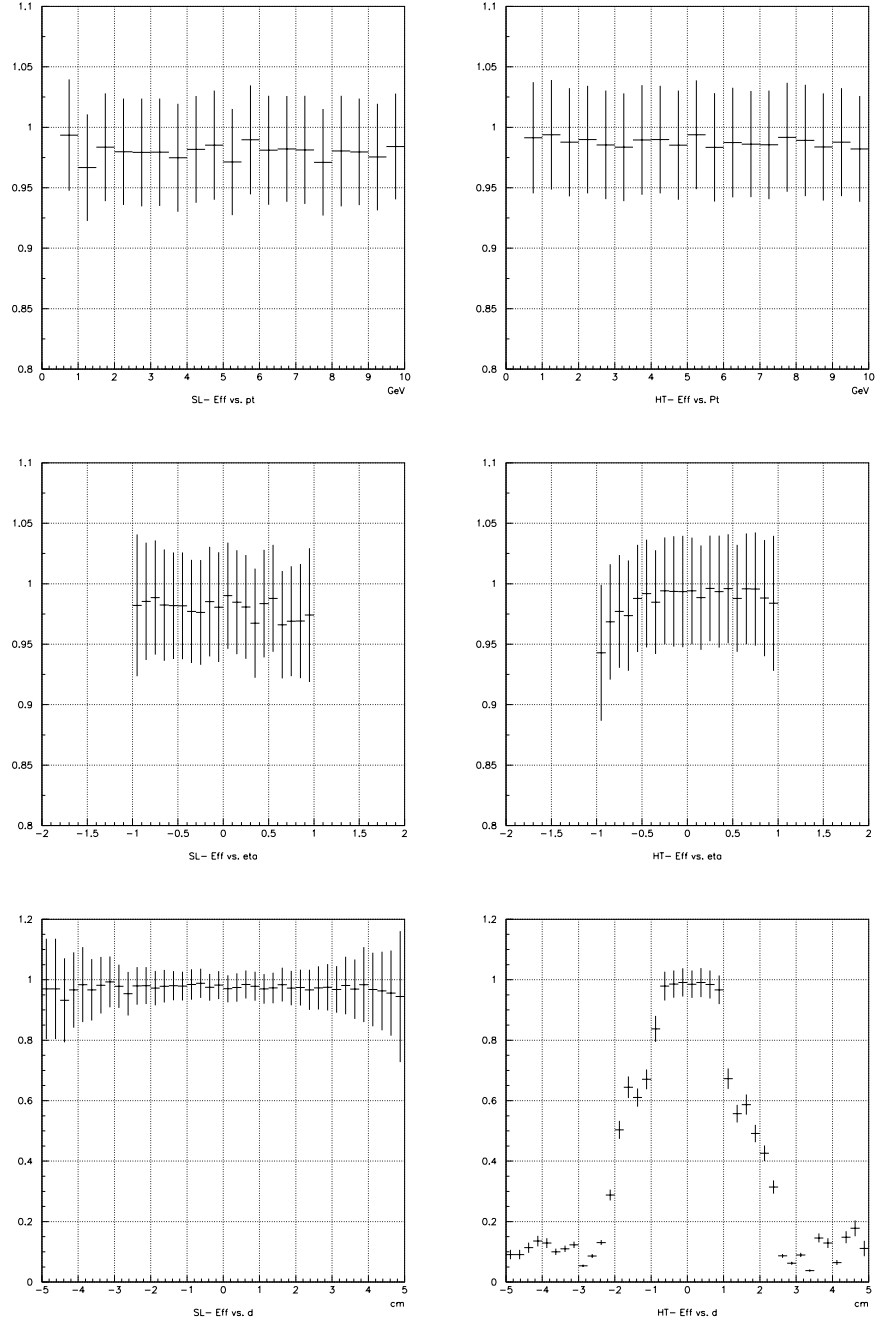
Figure 7.15: Fit-based efficiency on single track events, as a function of track parameters, for Segment Linking (left) and constrained Histogram Tracking (right). Efficiency as a function of p_t and η is computed on a sample of tracks with zero impact parameter; a different sample, containing tracks of varied impact parameter, is used for efficiency vs d.
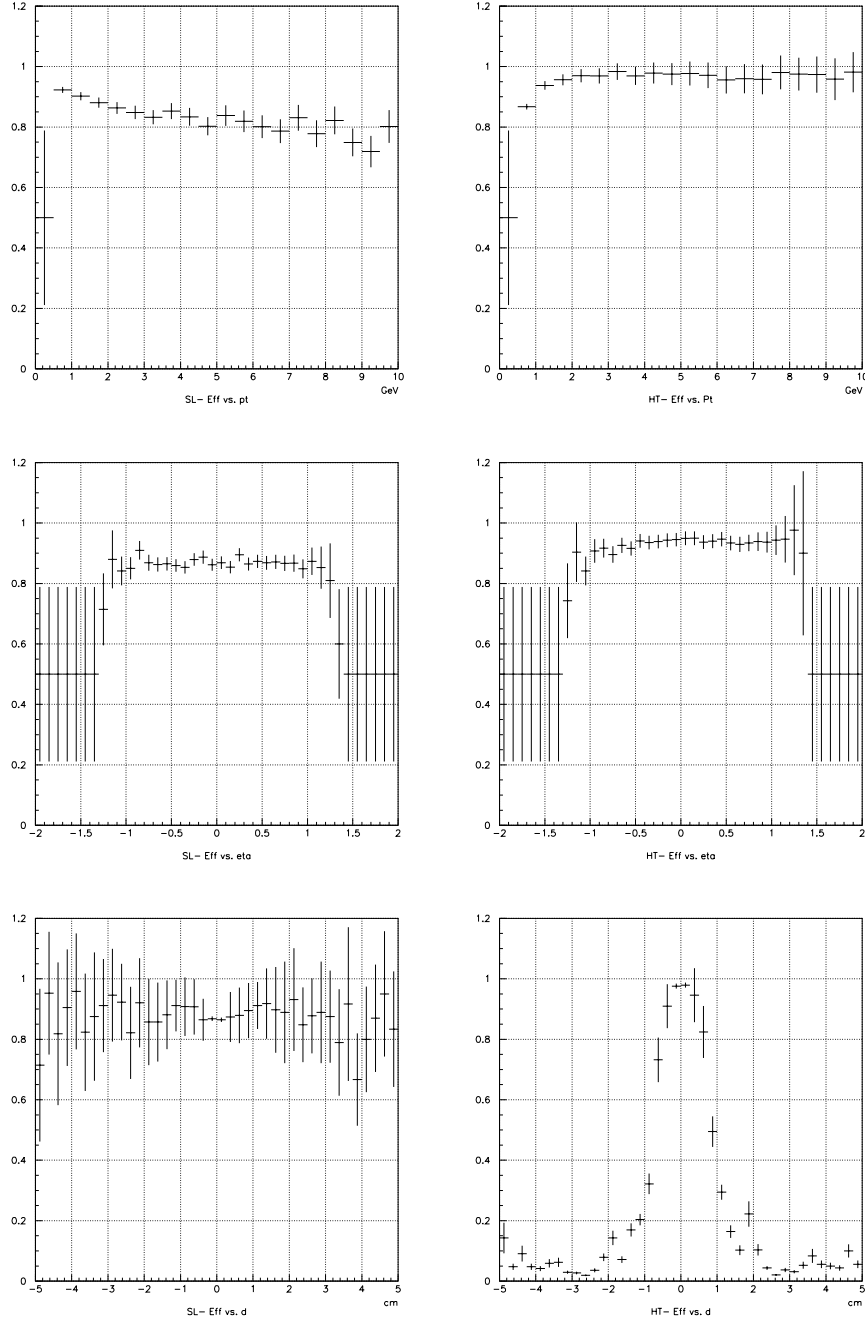
Figure 7.16: Fit-based efficiency on $t\bar{t}$ events, as a function of track parameters, for Segment Linking (left) and constrained Histogram Tracking (right).
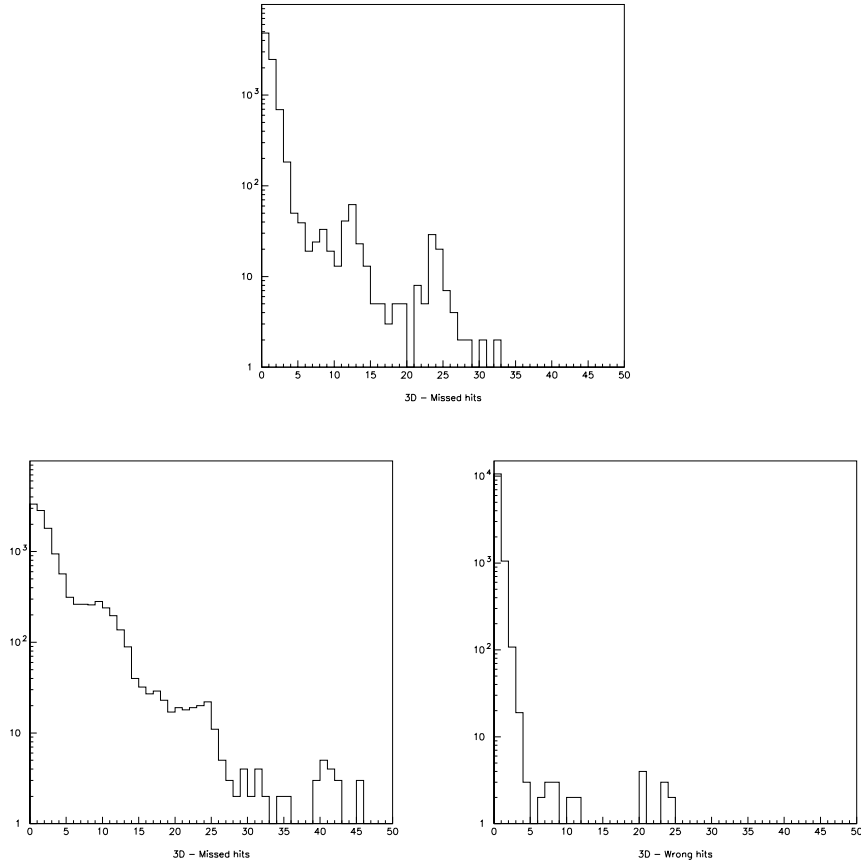
Figure 7.17: Missed stereo hits per track in COT stereo reconstruction on single track events (above); missed and wrong hits in $t\bar{t}$ events (bottom left and right). Tracks are required to have at least 40 clean stereo hits; axial reconstruction must have succeeded.

Chapter 8

Tracking in the silicon detectors

Tracking in the Central Outer Tracker begins with the "low-level" task of converting TDC signals into hit positions. In the same way, the first step of silicon tracking consists in transforming the amounts of charge deposited on each strip into hits.

After describing this preliminary step, this chapter shows briefly how the silicon detectors are used to determine the z coordinate of the primary vertex (or vertices); the rest of the chapter is devoted to the description and analysis of the silicon track reconstruction algorithms used at CDF II.

The general CDF II silicon tracking framework is presented first; the description of the specific algorithms (Stand-alone, Outside-In and Histogram) follows. The final section studies the performance of silicon pattern recognition, and reports the time cost and efficiency of the tracking phases.

8.1 Clustering

As was stated in section 3.4.3, when a charged particle crosses a biased silicon microstrip detector, an electron-hole cloud is generated along the path, and a small amount of electric charge is collected on one or more strips. The signal on each strip is then compared to the channel's pedestal, and — if the strip's charge is above threshold — it is read out. The raw output of the CDF II vertex detectors consists therefore in a list of active strips, together with the corresponding deposited charge. Before any track reconstruction is attempted, it is necessary to group adjacent strips into hits, and to calculate their position and accuracy; this is done in the **clustering** phase.

Clustering is performed separately on each half-ladder, as a series of consecutive steps. First of all, the clustering algorithm scans each side of the half-ladder, looking for a set of consecutive, above-threshold strips (a **super-**

cluster). Superclusters are split into clusters by looking for local minima in the profile, as shown in figure 8.1. The strip at the minimum (if any) is assigned to the closest peak's cluster; if the two closest peaks are at the same distance, the strip charge is equally split among them.
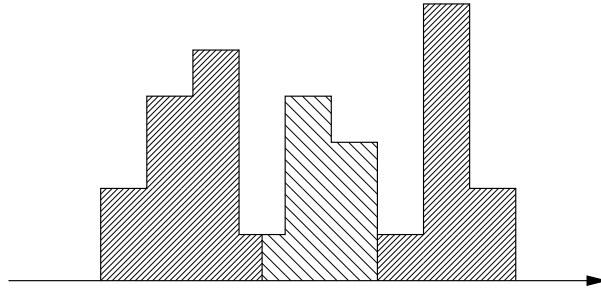
Figure 8.1: Example of silicon supercluster

Candidate clusters are then required to pass a set of quality cuts. These cuts currently consist in a maximum limit to the cluster length, and in a minimum cluster peak height (with the threshold being a function of cluster length). As a result of the first cut, the detector regions where the occupancy is extremely high are marked as "dirty"; they are not used to reconstruct hit positions, as the result would not be trustworthy. The second cut, on the other hand, helps in rejecting fake hits (which are due to noise fluctuations).

Studies are in progress on a more elaborate version of the clustering algorithm: the code should try to merge adjacent cluster candidates that do not pass the peak height cut, and to prune strips off the ends of very long clusters. The main worry about this approach is that the clustering algorithm must be proven to terminate within a reasonable amount of time — which is not a trivial issue if the strategy can both widen (by merging) and narrow (by pruning) the candidate clusters.

The final step consists in determining the charge centroid of each cluster, and in applying charge-deposition corrections. In an ideal detector, the centroid of each cluster would correspond to the intersection of a track with the plane of the wafer. In practice, the magnetic field affects the drift direction of charge bearers inside silicon in a non-negligible way, causing a shift in the centroid position.

8.2 Primary vertex

As was anticipated in section 7.4.3, as soon as stereo clusters have been reconstructed in the three 90° layers of SVX II, their positions are used to determine the approximate z positions of primary vertices. This position is then used to ease three-dimensional reconstruction of tracks.

The primary vertex finder begins by looping over the twelve azimuthal wedges, and by counting how many 90° hits were found on the inner (n_i) and on the outer (n_o) layer of that wedge. The number of possible combinations is $N = n_i n_o$, while the number of clean tracks with one z hit on both layers can be at most $S = \max(n_i, n_o)$. Wedges are then sorted in decreasing order of signal-to-noise (S/N) ratio; at this point, it is possible to discard the most noisy wedges.

Each pair of inner and outer hits determines a straight line uniquely; the algorithm finds its intersection with the middle layer, and checks whether the intersection is close enough to one of the middle layer hits. If this is true, the track is extrapolated to the beam axis; the intersection is then added to a histogram.

Once this task is done, the histogram profile is smoothed, adding to each bin a fraction of the contents in the adjacent bins. The algorithm then locates the global maximum, and clears the bins ranging from that maximum to the closest local minima. If the current maximum is high enough (compared both to an absolute threshold and to the previously found maxima), the current z is accepted as a primary vertex position; the algorithm then repeats the previous steps in order to find (if there are any) the other primary vertices of the event.

8.3 Silicon tracking framework

Once silicon hits have been found by the clustering algorithm, it is necessary to connect them to each other and to COT tracks, in order to reconstruct the event's set of charged particle tracks.

In the COT, which — apart from the reduced cell size — has a geometry closely resembling that of its Run I predecessor, the baseline tracking code was adapted from Run I, with a simple translation from FORTRAN into (mostly procedural) C++. On the contrary, the current silicon tracker — with its eight layers, seven of which double-sided — is substantially different from the four single-sided layers of Run I, and offers new potentialities, such as stand-alone silicon tracking. As a consequence, the Run II silicon tracking code has been written from scratch, making use of object-oriented techniques.

Due to the complexity of 3D tracking in a high-occupancy environment, and in order to allow easy comparison between competing algorithms, silicon tracking code is organized in interchangeable **Strategy** C++ classes. This object-oriented approach offers several advantages. All strategies have the same interface: they are given the current status of the detector as input — including partial results from previous strategies — and have to report found tracks in a standard way. It is therefore easy to apply different algorithms to the same data and to concatenate complementary strategies. Moreover, it is possible to decide at run-time which strategies are used and in which order, depending on the current situation.

Complete reconstruction of an event can take a long time. For this reason, several strategies support "regional tracking". Some "interesting" parts of the tracking volume are selected via simple criteria, such as matching a certain calorimeter tower; tracks are only reconstructed in that section of the detector. This can noticeably speed up the performance of the level 3 trigger, which uses the same reconstruction software as the offline analysis.

8.4 Stand-alone Tracking

A possible strategy for reconstruction of tracks in the silicon detector consists in finding triplets of aligned 3D hits, extrapolating them, and adding matching 3D hits on other layers. This strategy does not require any input from outside: tracking is performed completely inside the vertex detector.

As a first step, the algorithm builds three-dimensional hits from all the possible couples of intersecting axial and stereo strips on each wafer. When a small-angle stereo layer is crossed by a single track, and if there are no fake hits due to noise, then the 3D hit is unambiguous; multiple tracks with a small separation along φ can lead to ambiguous hit candidates, as shown in figure 8.2.

The situation is much worse on 90° layers, not only because the larger stereo angle leads to a greater number of ambiguous hits (as in the figure), but also because 90° strips on SVX II are multiplexed: each readout channel corresponds to two, three or four strips, depending on the layer.

Once a list of 3D hits is available, the Stand-alone silicon tracking algorithm searches it for triplets of aligned hits. Since ambiguous 3D hits are extremely more common on 90° layers, the triplet search is only performed on the four small-angle stereo layers.

In order to perform this search, the algorithm loops over the couples of unused 3D hits in the "inner" and the "outer" layer. For each hit pair, it draws a straight line in the rz plane. The intersection of this line with the
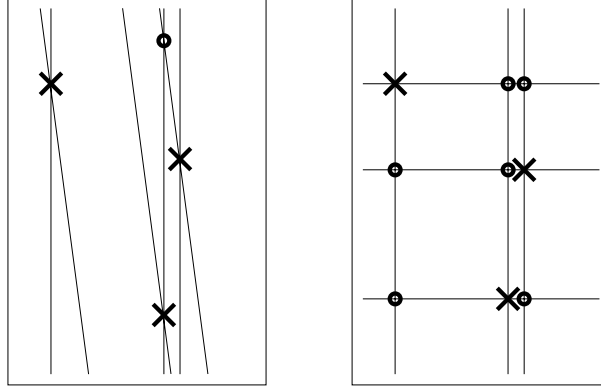
Figure 8.2: Ambiguity in double-sided silicon detectors, with a small stereo angle (left) or a 90° angle (right, without multiplexing). True hits are marked with crosses, spurious hits with circles.

z axis is required to be compatible with one of the previously found primary vertex z bins; the line slope is also examined, to see if the current track seed could intersect the other silicon layers.

At this point the "middle" layer is examined; each of its hits is used to build a helix together with the current inner and outer hits. This helix must exist (within the error), must pass some configurable "interest" cuts (for example, the track momentum can be required to exceed a threshold), and its extrapolation to the beam axis must be close to a primary vertex.

Once the list of candidates is ready, each triplet is compared to all the others; if two triplets have enough hits in common, and the remaining hits are compatible, the triplets are merged and the helix is recalculated.

Each candidate is then extrapolated to all the other silicon layers; the algorithm scans the proximity of each candidate-wafer intersection for hits to add to the track. If there is only one such hit, it is added; if there are two or more, the candidate is cloned, and a different hit is attached to each candidate. Axial hits are scanned first, then small-angle stereo hits, and finally z hits. Candidates are required to contain at least five axial hits, three small-angle ones (from the seed triplet) and two 90° hits; those that do not are discarded.

Full helix fits (which take multiple scattering in account) are then performed on all the candidates. The fit results are compared among each group of candidate clones. A figure of merit is determined summing the fit χ^2 to a configurable missed-layer penalty; the best candidate is kept, the others are rejected. As a final step, any remaining candidates are compared to each

other in order to resolve conflicts in hit attribution.

The main disadvantage of Stand-alone silicon tracking is its extreme dependance on detector occupancy: the algorithm needs to perform three nested loops on the seed layers, scanning a number of candidate 3D hits that is itself proportional to the square of the layer's occupancy. In events with many tracks, this means having to examine a very large number of candidates, spending a very long time on computations, and having a high chance of associating uncorrelated hits into fake tracks. On the other hand, since it does not require outside information, Stand-alone Tracking is the only way to use the CDF II silicon detectors to their full η range, beyond the $|\eta| < 1$ acceptance limit of the COT.

For these reasons, silicon tracking in CDF II begins with one or more cleanup steps, which combine COT and silicon information to reconstruct very reliable tracks in the central, $|\eta| < 1$, region; once the hits from these tracks have been marked as used, Stand-alone silicon tracking can be used to find forward $(1 < |\eta| < 2)$ tracks in a less crowded environment.

8.5   Outside-In Tracking

One algorithm that will be used at CDF II to extend COT tracks into the silicon is called "Outside-In Tracking". In order to avoid the previously described ambiguity in 3D hits, this algorithm performs two subsequent reconstruction steps: it first attaches axial silicon hits to COT tracks, and then performs pattern recognition on small-angle and 90° stereo strips. The number of hits to be examined decreases thus from $O(n^2)$ to $O(2n)$; this reduction is made possible by the use of an external seed.

During each of the Outside-In phases, each seed track is extrapolated towards the silicon, until it intersects (or passes sufficiently close to) a wafer. The algorithm then scans the wafer for hits of the proper kind (axial or stereo). Supposing there are N candidate hits, N copies of the current seed are created, and a different hit is attached to each; the resulting tracks are fitted to helices via a progressive (Kalman) filter [65]. At the same time, the original seed is kept unchanged in the candidates list.

Each of the resulting helices is then extrapolated to the next silicon wafer; once again, each candidate is replicated in a certain number of copies, each of which is attached to a different hit on the new layer and re-fitted. This process is repeated until the extrapolations reach the innermost layer; each seed track generates thus a search graph like the simplified one shown in figure 8.3. Each path from the root (the seed track) to one of the leaf nodes corresponds to a different choice of the hits to be attached to the current
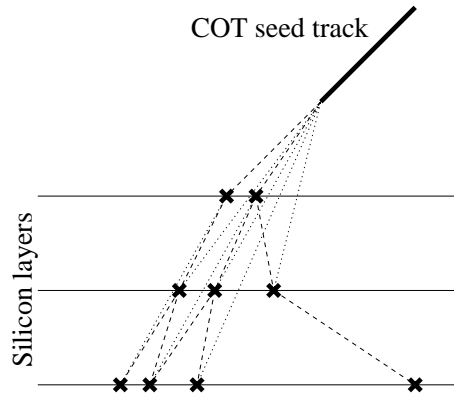
Figure 8.3: Simplified Outside-In search graph

track, and to a different helix fit. In the end, the candidate with the highest number of hits and the lowest χ^2 per degree of freedom is kept as a found track; all the other possibilities are discarded.

Outside-In is the baseline tracking algorithm for axial reconstruction in the silicon; the code is being extended to work on stereo silicon hits as well.

8.6 Histogram Tracking

While being capable of a quite good performance, Outside-In silicon tracking suffers from one major drawback: speed. For each seed, the algorithm needs to examine a potentially large number of candidates, which increases with a high power of occupancy. Moreover, the inner loop is slow, since it performs a helix fit on each candidate, also taking into account the actual material traversed by the track.

Outside-In could be a good choice for offline analysis, where computation time is not a critical issue; in online reconstruction, however, speed is an essential factor, and a faster algorithm can improve performance noticeably on dense events. Outside-In is slow because of its combinatory behaviour, and because of the slow inner loop. The Histogram Tracking algorithm, described in section 7.5, does not suffer from either of these problems: applying it to the silicon detector is a natural alternative.

Like Outside-In Tracking, Histogram Tracking in the silicon consists of two consecutive steps: it begins by reconstructing tracks in the transverse plane, and then attaches stereo hits to them. Beside the reasons given above, this separation in consecutive phases is made necessary by the 1-dimensionality of the strategy's histograms.

8.6.1   R-φ Histogram Tracking

The first part of Histogram Tracking in the silicon consists in attaching axial hits to each fully fitted, 3D COT track.

The process begins, as Outside-In Tracking, by extrapolating the seed track into the silicon detectors; only the strips which lie close enough to the extrapolated track are further examined. Usage of COT 3D information greatly reduces the number of candidate hits, as it makes it possible to select one, or at most two, out of the six half-ladders that make up each azimuthal wedge of each layer.

Each silicon hit within the search region is then used, together with the seed data, to add a contribution to the current layer's histogram. Due to the same performance issues that were explained in section 7.5.1 when describing Histogram Tracking in the COT, the shape of each single-hit contribution to $\log \mathcal{L}(\mathcal{D})$ is approximated with a parabola. The algorithm, therefore, needs to compute three quantities for each hit: the "best" value of the histogram abscissa $\mathcal{D}$, the value of the likelihood at the peak, and its second derivative with respect to $\mathcal{D}$. Figure 8.4 provides a schematic illustration of this process.
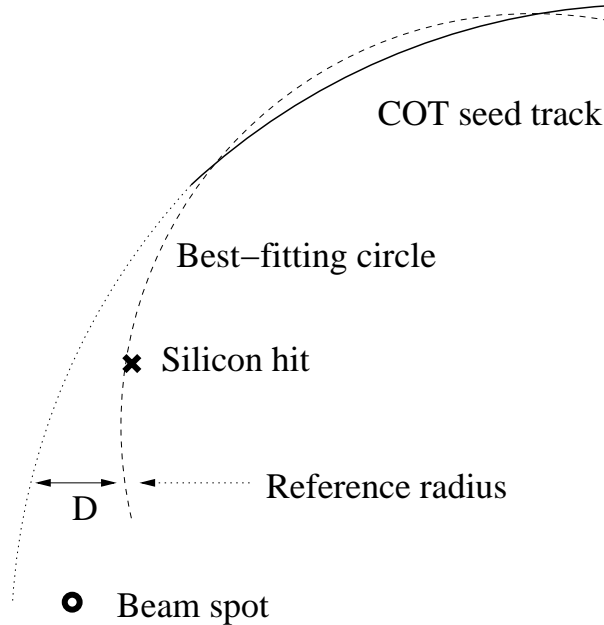
Figure 8.4: R-φ Histogram Tracking: single hit contribution

The algorithm determines the circle on the transverse plane that best fits the COT track and the current hit, taking in consideration their respective accuracy. This "hit circle" is then compared to the "seed circle" defined

by the seed track: the histogram abscissa $\mathcal{D}$ is defined as the azimuthal arc length between the two circles (see figure 8.4), at an arbitrary value of r (the "reference radius" r_R). The calculation is explained in detail hereafter.

Let $\vec{\alpha} = (c, d, \varphi_0)$ be the set of parameters describing a circle in the transverse plane; let $\vec{\alpha}_0$ correspond to the COT seed track, and $\vec{\alpha}^*$ to the best-fitting circle - the one that minimizes

$$\chi^2 = (\vec{\alpha} - \vec{\alpha}_0)^T \Lambda (\vec{\alpha} - \vec{\alpha}_0) + \left(\frac{\delta(hit, \vec{\alpha})}{\sigma_{hit}} \right)^2 \tag{8.1}$$

where Λ is the inverse of the seed's covariance matrix, restricted to the transverse plane, σ_{hit} is the uncertainty on the hit (as provided by the clustering algorithm), and $\delta(hit, \vec{\alpha})$ is the distance between the current hit and the α-circle's intersection with the hit's wafer.

Since the circle equation

$$\varphi(r; c, d, \varphi_0) = \varphi_0 + \arcsin\left(\frac{cr + d(1 + cd)/r}{1 + 2cd} \right) \tag{8.2}$$

is non-linear, the χ^2 minimization problem is non-linear itself. In order not to slow down computation excessively, it is necessary to use a linear approximation. First of all, the algorithm assumes that the silicon wafer is a cylinder of radius r_{int}, where r_{int} is the radial coordinate of the intersection of the seed track with the wafer's plane. The hit–intersection distance is also expected to be small. As a result of these approximations,

$$\delta(hit, \vec{\alpha}) \simeq r_{int} (\varphi_{hit} - \varphi(r_{int}; \vec{\alpha})) \tag{8.3}$$

It is then necessary to linearize equation 8.3 around the seed circle $\vec{\alpha}_0$:

$$\varphi(r_{int}; \vec{\alpha}) \approx \varphi(r_{int}; \vec{\alpha}_0) + \left. \frac{\partial \vec{\varphi}}{\partial \vec{\alpha}} \right|_{\vec{\alpha}_0} \cdot (\vec{\alpha} - \vec{\alpha}_0) \equiv \bar{\varphi} + \vec{v} \cdot (\vec{\alpha} - \vec{\alpha}_0) \tag{8.4}$$

where

$$\bar{\varphi} \equiv \varphi(r_{int}; \vec{\alpha}_0) \tag{8.5}$$

$$\vec{v} \equiv \left. \frac{\partial \vec{\varphi}}{\partial \vec{\alpha}} \right|_{\vec{\alpha}_0} \equiv \left. \begin{pmatrix} \frac{\partial \varphi}{\partial c} \\[2mm] \frac{\partial \varphi}{\partial d} \\[2mm] \frac{\partial \varphi}{\partial \varphi_0} \end{pmatrix} \right|_{(\bar{c}, \bar{d}, \bar{\varphi}_0)} = \left. \begin{pmatrix} \frac{1}{\sqrt{1-F^2}} \frac{r_{int} - \frac{d^2}{r_{int}}}{(1+2cd)^2} \\[2mm] \frac{1}{\sqrt{1-F^2}} \frac{1}{r_{int}} \frac{1+2cd+2(cd)^2 - 2(cr_{int})^2}{(1+2cd)^2} \\[2mm] 1 \end{pmatrix} \right|_{(\bar{c}, \bar{d}, \bar{\varphi}_0)} \tag{8.6}$$

$$F = \frac{c r_{int} + d(1 + cd)/r_{int}}{1 + 2cd} \tag{8.7}$$

With this approximation, the linear equations that express the χ^2 minimization are

$$0 = \frac{\partial \chi^2}{\partial \vec{\alpha}} = 2 \left[\Lambda \cdot \Delta \vec{\alpha} - \frac{r_{int}^2}{\sigma_{hit}^2} \left(\varphi_{hit} - \bar{\varphi} - \vec{v} \cdot \Delta \vec{\alpha} \right) \vec{v} \right] \tag{8.8}$$

where $\Delta \vec{\alpha} = \vec{\alpha}^* - \vec{\alpha}_0$. The solution of (8.8) is

$$\vec{\alpha}^* = \vec{\alpha}_0 + (\Lambda + M)^{-1} \cdot \vec{w} \tag{8.9}$$

where

$$M \equiv \left(\frac{r_{int}}{\sigma_{hit}} \right)^2 \vec{v} \otimes \vec{v} \tag{8.10}$$

is a 3×3 symmetric matrix, and

$$\vec{w} \equiv \left(\frac{r_{int}}{\sigma_{hit}} \right)^2 \left(\varphi_{hit} - \bar{\varphi} \right) \vec{v} \tag{8.11}$$

The 3×3 covariance matrix associated to the solution (8.9) is

$$\mathcal{C} = K^{-1} = (\Lambda + M)^{-1} \tag{8.12}$$

as can be easily obtained from (8.8) given that

$$K = \frac{1}{2} \frac{\partial^2 \chi^2}{\partial \vec{\alpha} \partial \vec{\alpha}^t} \tag{8.13}$$

Given $\vec{\alpha}^*$ and $\vec{\alpha}_0$, it is finally possible to determine the value of $\mathcal{D}$ that maximizes the likelihood of the track passing through the hit. As was stated before, $\mathcal{D}$ is the azimuthal arc length between the hit circle and the seed circle, computed at the (constant) radial coordinate r_R:

$$\mathcal{D}_{best} = r_R \left(\varphi(r_R; \vec{\alpha}^*) - \varphi(r_R; \vec{\alpha}_0) \right) \tag{8.14}$$

The value of the χ^2 minimum, that is the value of (8.1) evaluated at the solution (8.9), is

$$\chi^2_{min} = (K\vec{w})^t \, \Lambda \, K\vec{w} + \left(\frac{r_{int}}{\sigma_{hit}} \right)^2 \left(\varphi_{hit} - \bar{\varphi} - \vec{v} \cdot K\vec{w} \right)^2 \tag{8.15}$$

As in the COT case, it is possible to rescale the likelihood, so that $\log \mathcal{L} = \chi^2_{CUT} - \chi^2(\vec{\alpha})$, where χ^2_{CUT} is a constant, chosen so that the log-likelihood of the track not having any hit on the current layer is zero. As a consequence, at $\mathcal{D} = \mathcal{D}_{best}$, $\log \mathcal{L}$ assumes the peak value

$$\mathcal{L}_{best} = \chi^2_{CUT} - \chi^2_{min} \tag{8.16}$$

In the COT, the hit circle is perfectly determined: the number of constraints is equal to the number of degrees of freedom for a circle, as explained in Section 7.5.3, and the peak height is the same for all hit candidates. In the silicon, the circle is overconstrained: for this reason, $\mathcal{L}_{best}$ can assume different values for different hits.

To complete the description of a single hit contribution, the algorithm propagates the error from the telescope and the hit onto $\mathcal{D}_{best}$, thus determining the width of the parabola.

Once all the hits have been examined, and their contributions superimposed to the corresponding layer histograms, the tracking algorithm resorts to the same procedure described in 7.5.1: it sums the histograms, finds the global maximum in the overall likelihood, and selects the hits which contributed to this maximum.

8.6.2 Selection of hits

Silicon detectors and drift chambers differ both in the principles of operation and in their geometry; this entails some differences in the process of hit selection. In the COT, it was sufficient to select the best hit on each layer; in the silicon, there are reasons that make it necessary to use a more complex procedure.

On one hand, in order to maximize the geometric acceptance of the detector, it was necessary to introduce an overlap region between adjacent φ wedges. Without such an overlap, tracks with certain values of φ and curvature could slip through a layer undetected[1]. As a consequence of the overlap, though, the same track can leave multiple hits in adjacent wedges of the same layer. Tracking algorithms, therefore, must be able to select multiple hits in the overlap regions, provided the hits are on different φ wedges.

On the other hand, hit selection in the silicon is a much more sensitive process than it is in the COT. First of all, the number of layers is smaller, and the hit resolution is better; single silicon hits are therefore more important than single COT hits. Moreover, due to their closeness to the interaction point, silicon strips have an extreme importance in determining the track's d and z_0 parameters — and hence the positions of vertices. A single wrong hit on an inner silicon layer can easily fake the presence of a secondary vertex in the event, thus reducing the purity of b-physics samples. Finally, choosing

---

[1]The geometric acceptance of the silicon tracker is not 100%, due to the presence of "blind" regions at the z boundary between subsequent wafers. Contrarily to the φ cracks, the z ones cannot be closed by introducing an overlap region — not without causing worse problems in terms of detector alignment and/or of increased multiple scattering.

the wrong axial hit on a small-angle stereo layer means an automatic failure
in the subsequent 3D tracking step.

For these reasons, selection of silicon hits needs to be performed with a
more accurate algorithm: selecting the hits that contributed to the global
histogram peak is only the first step.

Histogram Tracking relies on the best-$\mathcal{D}$ values of the "good" hits to be
very close to each other. In the ideal case, any differences should be due to
the finite resolution of candidate hits, and should be little correlated to each
other. In practice, however, the error on the telescope parameters propagates
with different intensity to the different layers; this leads to a linear correlation
between r and the $\mathcal{D}$ shift. Therefore, good candidates tend to form a line
in the $r\mathcal{D}$ plane that is not perfectly vertical, but has a finite slope.

Unfortunately, it is not possible to eliminate this slope completely by
mathematical transformations: the error on the seed is distributed over sev-
eral degrees of freedom, and is not perfectly correlated. A proper choice of $\mathcal{D}$
can eliminate the slope due to a **single** degree of freedom: the telescope error
along that direction of parameter space is transformed to an equal $\mathcal{D}$-shift
on all the layers. Errors on the other degrees of freedom of the seed lead to
inhomogeneous $\mathcal{D}$ shifts, which are approximately a linear function of r.

The r-$\mathcal{D}$ slope is a drawback in the "general" part of Histogram Tracking.
In order to form a clear maximum in the global histogram, the contributions
of the various "good" hits must overlap; if the contributions are not all
centered on the same value of $\mathcal{D}$, it is possible that a spurious hit might
match the global maximum better than the good hit, as exemplified in figure
8.5.
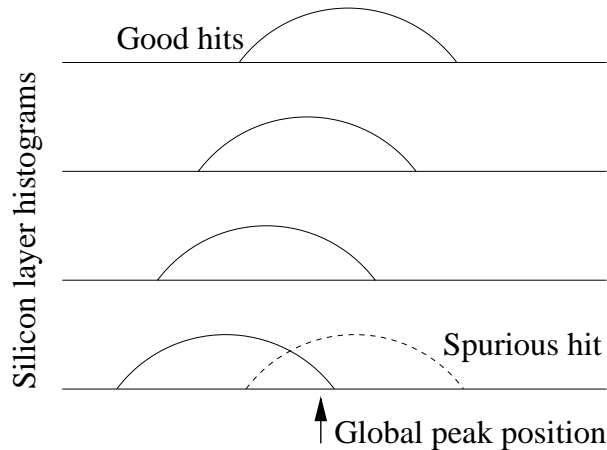
Figure 8.5: Sloping alignment in the rD plane

In order to avoid this problem, it is possible to loosen the $\mathcal{D}$ cut, possibly selecting mutually incompatible hit candidates on the same layer, and then to filter the candidates by requiring them to lie roughly on the same sloping line. For this purpose, the candidate hits are divided in three groups, in descending order of trustability:

- Unique candidates on their layer

- Pairs of candidates on the same layer, but on adjacent wedges (possibly due to the overlap)

- Ambiguous hits (multiple hits on the same layer and wedge)

The algorithm tries to fit a straight line through the most reliable points: unique candidates, if there are enough; unique and overlap candidates, as a second choice. If there are not enough unambiguous candidates, hit selection falls back to a simpler algorithm — discard the hopelessly ambiguous layers, select the candidates which lie closest to the peak, and hope for the best.

If the algorithm manages to fit a straight line through trusted points, it proceeds to resolve ambiguities, by selecting on each ambiguous layer the candidate[2] that lies closest to the extrapolation of the fitted line.

At this point, all the "good" hits should have been selected; however, some spurious ones might have survived. In order to remove them, each hit candidate is compared to the straight line fit of the other candidates. The algorithm finds the worst matching hit; if its residual is too large, the hit is removed and the cleanup step starts anew.

While this selection algorithm performs well in the most common cases — either an ambiguous set of candidate hits, or an unambiguous set of hits some of which are spurious — it proves vulnerable when the set of candidates contains both ambiguities and unambiguous fake hits. The unambiguous fake hit might shift the "trusted" line so much that, when it is used to resolve ambiguities, more fake hits will be selected; given this situation, the clean-up step might remove the remaining good hits and worsen the situation. To avoid this, the initial selection of candidates must be tuned to avoid the contemporary presence of ambiguities and of isolated fakes — even at the cost of a slight decrease in efficiency.

8.6.3 R-Z Histogram Tracking

Once axial hits have been attached to a COT track, either by histogramming or with the Outside-In algorithm, Histogram Tracking can be used to

---

[2]Or the pair of overlapping candidates, provided they belong to different azimuthal wedges.

complete reconstruction on the RZ plane.

As in Histogram R-φ Tracking, the first step consists in selecting a set of silicon strips that could contain hits belonging to the track. On the three 90° layers, the fit parameters of the seed track determine a search region directly. On the small-angle stereo layers, instead, the algorithm searches the half-ladders on which the seed track has a previously found axial hit: half-ladders without an axial hit are not used. The previous reconstruction step guarantees to select at most one axial hit per wafer; therefore, there is no combinatory ambiguity. By comparing the axial hit position to each compatible stereo hit, the algorithm determines a set of possible z coordinates; those close enough to the seed track are kept for further use.

The next step consists, as usual, in building the layer histograms, by superimposing parabolic contributions from each of the layer's hit candidates. This is done by approximating the rz projection of the trajectory with a straight line that fits both the seed parameters and the hit position, each within the respective errors, as shown in figure 8.6. The distance along z between the "seed line" and the "hit line", at a radial coordinate r_R, is used as the histogram abscissa $\mathcal{D}$. Each silicon wafer is approximated with a cylindrical surface, whose radius is equal to r_{int} — the radial coordinate of the intersection of the seed track with the wafer itself.
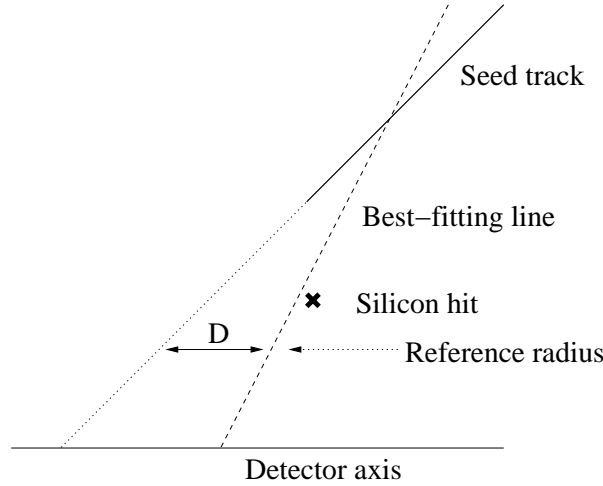
Figure 8.6: R-Z Histogram Tracking: single hit contribution

As before, the "hit line" is determined by χ^2 minimization. Let z_{hit} be the position of the current hit, and σ_{hit} its accuracy[3]; let $\cot\theta_0$ and z_{00} be

---

[3]The z accuracy of 90° hits is much better than that of 3D, small-angle stereo hits. This is taken into account.

the parameters of the seed track, whose 2D covariance matrix in the rz plane is Λ^{-1}; let $\cot\theta$ and z_0 describe a generic line in the rz plane, while $\cot\theta^*$ and z_0^* are the parameters of the best line — the one that minimizes

$$\chi^2 = (\cot\theta - \cot\theta_0, z_0 - z_{00})\,\Lambda\begin{pmatrix}\cot\theta - \cot\theta_0 \\ z_0 - z_{00}\end{pmatrix} + \left(\frac{z_{hit} - z_0 - r_{int}\cot\theta}{\sigma_{hit}}\right)^2$$

(8.17)

By inverting the covariance matrix,

$$\Lambda^{-1} = \begin{pmatrix} \sigma^2_{\cot\cot} & \mathrm{COV}_{z_0\cot} \\ \mathrm{COV}_{z_0\cot} & \sigma^2_{z_0} \end{pmatrix}$$

(8.18)

one obtains

$$\begin{aligned}
\Lambda &= \frac{1}{\sigma^2_{\cot}\sigma^2_{z_0} - \mathrm{COV}^2_{z_0\cot}}\begin{pmatrix} \sigma^2_{z_0} & -\mathrm{COV}_{z_0\cot} \\ -\mathrm{COV}_{z_0\cot} & \sigma^2_{\cot} \end{pmatrix} \\
&= \frac{1}{\sigma^2_{\cot}\sigma^2_{z_0}(1-\rho^2)}\begin{pmatrix} \sigma^2_{z_0} & -\mathrm{COV}_{z_0\cot} \\ -\mathrm{COV}_{z_0\cot} & \sigma^2_{\cot} \end{pmatrix}
\end{aligned}$$

(8.19)

where

$$\rho \equiv \frac{\mathrm{COV}_{z_0\cot}}{\sigma_{\cot}\sigma_{z_0}}$$

(8.20)

is the correlation coefficient. It is also useful to define the adimensional coefficient $\beta \equiv 1/(1-\rho^2)$.

By requiring the partial derivatives of χ^2 with respect to $\cot\theta$ and z_0 to be null, one obtains the system of linear equations

$$\begin{aligned}
\left(\frac{1}{\sigma^2_{hit}} + \frac{\beta}{\sigma^2_{z_0}}\right)z_0^* + \left(\frac{r_{hit}}{\sigma^2_{hit}} - \frac{\rho\beta}{\sigma_{\cot}\sigma_{z_0}}\right)\cot\theta^* &= \frac{z_{hit}}{\sigma^2_{hit}} + \frac{\beta}{\sigma^2_{z_0}}z_{00} - \frac{\rho\beta}{\sigma_{\cot}\sigma_{z_0}}\cot\theta_0 \\
\left(\frac{r_{hit}}{\sigma^2_{hit}} - \frac{\rho\beta}{\sigma_{\cot}\sigma_{z_0}}\right)z_0^* + \left(\frac{r^2_{hit}}{\sigma^2 hit} + \frac{\beta}{\sigma^2_{\cot}}\right)\cot\theta^* &= \frac{z_{hit}r_{int}}{\sigma^2_{hit}} - \frac{\rho\beta}{\sigma_{\cot}\sigma_{z_0}}z_{00} + \frac{\beta}{\sigma^2_{\cot}}\cot\theta_0
\end{aligned}$$

(8.21)

which is easily solved, as any linear system of the form

$$a_1 z_0^* + b_1\cot\theta^* = c_1$$
$$a_2 z_0^* + b_2\cot\theta^* = c_2$$

by the formulae

$$z_0^* = \frac{c_1 b_2 - b_1 c_2}{a_1 b_2 - b_1 a_2}$$
$$\cot\theta^* = \frac{a_1 c_2 - c_1 a_2}{a_1 b_2 - b_1 a_2}$$

(8.22)

It is worth noticing that all the coefficients of the system 8.21 are the sum of a term, proportional to $1/\sigma_{hit}^2$, that only depends on the current hit's parameters, and of another term which depends only on the seed track. This fact is exploited to speed up reconstruction.

Once the "hit line" has been determined, it can be used to determine the parameters of the parabolic contribution to the current layer's histogram. The peak is placed at

$$\mathcal{D}_{best} = \left(z_0^* - z_{00}\right) + r_R(\cot\theta^* - \cot\theta_0) \qquad (8.23)$$

at a height

$$\log\mathcal{L}_{best} = \chi_{CUT}^2 - \chi_{min}^2 \qquad (8.24)$$

where χ_{CUT}^2 is, as in r-φ tracking, a predetermined constant, and

$$
\begin{aligned}
\chi_{min}^2 \;=\; & \beta\left[\left(\frac{z_0^* - z_{00}}{\sigma_{z_0}}\right)^2 + \left(\frac{\cot\theta^* - \cot\theta_0}{\sigma_{\cot}}\right)^2 - 2\rho\frac{z_0^* - z_{00}}{\sigma_{z_0}}\frac{\cot\theta^* - \cot\theta_0}{\sigma_{\cot}}\right] \\
& + \left(\frac{z_{hit} - z_0^* - r_{int}\cot\theta^*}{\sigma_{hit}}\right)^2 \qquad (8.25)
\end{aligned}
$$

Lastly, the algorithm computes the second derivative of $\log\mathcal{L}$ at $\mathcal{D}_{best}$. In the current code, an approximated result is used:

$$\left.\frac{\partial^2\chi^2}{\partial\mathcal{D}^2}\right|_{\mathcal{D}_{best}} \simeq \frac{1}{\sigma_{hit}^2} + \frac{\beta}{\sigma_{z_0}^2} \qquad (8.26)$$

Once again, when all the hits in the search region have been examined, the algorithm superimposes their contributions to the corresponding layer's histogram; it sums the histograms, seeks the global peak, and selects a set of actual hits with the procedure described in 8.6.2.
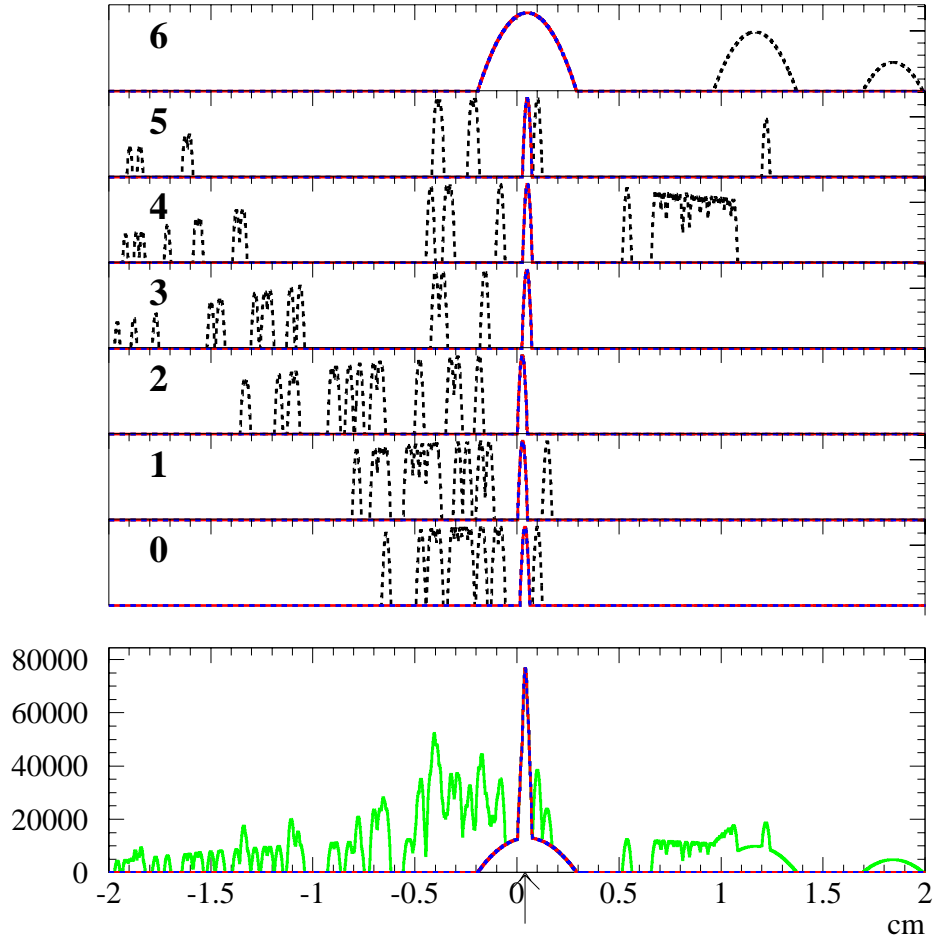
Figure 8.7: Example of Histogram Tracking on axial silicon hits. In the upper part of the figure, black dashed lines correspond to hits not belonging to the track; blue–red lines correspond to the hits correctly associated to the track. In the lower part, the green line is the sum histogram; the blue–red line is the contribution from correct hits.
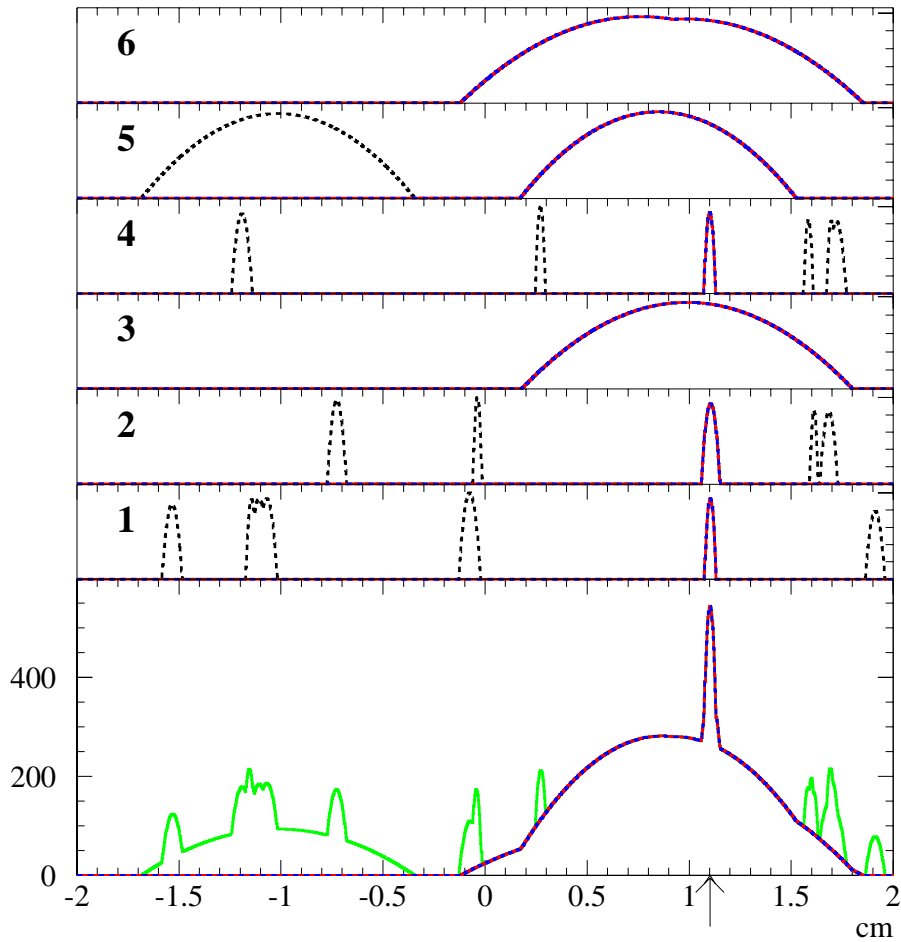
Figure 8.8: Example of Histogram Tracking on stereo silicon hits. In the upper part of the figure, black dashed lines correspond to hits not belonging to the track; blue–red lines correspond to the hits correctly associated to the track. In the lower part, the green line is the sum histogram; the blue–red line is the contribution from correct hits.

8.7 Performance

As was done for COT track reconstruction, the silicon tracking algorithms discussed above were tested on single-track and $t\bar{t}$ samples. Segment-Linking COT tracks were used to seed both Outside-In and Histogram Tracking[4]. Stand-alone silicon tracking was not taken in consideration: its performance depends too heavily on how well the previous reconstruction steps removed hits belonging to $|\eta| < 1$ tracks.

Table 8.1 lists the time consumption of the various strategies, in the same operating conditions as in section 7.6. The speed advantage of Histogram Tracking in the silicon is evident: with the default (and further optimizable) histogram binning, the actual pattern finding process is three times faster than Outside-In, while the entire reconstruction (clustering and tracking) is more than twice as fast. Silicon Histogram Tracking therefore fulfills its design goal of high speed.

Operation	Avg time per $t\bar{t}$ event (ms)
Clustering	670 ± 20
Axial Outside-In	3100 ± 200
Axial Histogram	
— 2000 bins (default)	1070 ± 60
— N bins	$\simeq 550 + 0.26N$
Stereo Outside-In	Code under development
Stereo Histogram	
— 2000 bins (default)	680 ± 70
— N bins	$\simeq 500 + 0.09N$

Table 8.1: Silicon reconstruction: CPU timing

The efficiency of each step in silicon tracking is, as usual, the ratio between "found" and "findable" tracks. As anticipated in Chapter 6, "found" tracks can be defined in two complementary ways: hit-based and fit-based. The former relies on the hit content of the tracks: tracking is declared successful if a findable track has

- No wrong hits;

- At most one missed layer;

---

[4]This choice is due to technical reasons: the seeds need to be 3D COT tracks, and 3D support has been added to Histogram Tracking only recently.

- Found hits on at least two 90° layers (for 3D tracking only).

The last requirement is necessary because, to determine both the slope and
z intercept of a track accurately, the track must contain at least two high-
precision rz measurements; COT and small-angle stereo hits do not provide
sufficient information.

Fit-based efficiency, on the other hand, completely overlooks the hit con-
tent of the track; if a "wrong" hit is perfectly aligned to the "good" ones,
then it can be safely accepted. Tracks are instead required to match their
ideal fit parameters, within 5σ of their ideal asymptotic resolution on the
most relevant parameters. Table 6.3 lists the cuts that were actually applied
to findable tracks.

"Findable" tracks are defined in the same way in both approaches. They
must:

- Have been correctly reconstructed in the previous tracking phases;

- Contain at least three clean axial layers (axial tracking), or (3D track-
  ing) at least two clean 90° layers and one further clean (90° or small-
  angle) stereo layer.

Hit– and fit-based efficiencies are shown in the subsequent figures, as
a function of the track's transverse momentum and pseudorapidity. The
average efficiencies for $p_t > 2$ GeV tracks are summarized in table 8.2.

Efficiency (%), $p_t > 2$ GeV	single tracks ($d = 0$)	$t\bar{t}$ events
Axial, Outside-In		
— Hit-based	99 ± 1	94 ± 1
— Fit-based	99.3 ± 1	96 ± 1
Axial, Histogram		
— Hit-based	99.7 ± 1	89 ± 1
— Fit-based	99.8 ± 1	95 ± 1
Stereo, Histogram		
— Hit-based	99.6 ± 1	93 ± 1.5
— Fit-based	99.5 ± 1	93 ± 1.5

Table 8.2: Global tracking efficiency in the silicon

In most cases, the two definitions of efficiency lead to matching results.
The only major disagreement can be found in axial Histogram Tracking on $t\bar{t}$

events, where fit-based efficiency is 95% (matching the outcome of Outside-In), but hit-based efficiency is only 89%. This is due to the fact that Histogram Tracking studies the correlation between hit positions only approximately, while Outside-In — with its slower progressive fitting — is able to take better decisions in borderline cases. Nevertheless, as the fit-based efficiency implies, wrong and missed hits selected by Histogram Tracking are never too far away from the real track.

In conclusion, when applied to the silicon vertex detectors of CDF II, the Histogram Tracking algorithm proved to reconstruct tracks quickly and with a fit-based efficiency comparable to the Outside-In procedure. Therefore, it is well-suited to be used as a component of the level 3 trigger. Outside-In Tracking, on the other hand, provides better efficiency on low-momentum tracks, and is more accurate in borderline decisions. It is probably the best choice in absence of strict time restrictions (such as in the offline analysis of relatively small samples). The Outside-In algorithm could also be used as a refinement step in the trigger: once Histogram Tracking has reconstructed the event, the (few) tracks directly involved in the trigger decision could be examined again with the more accurate algorithm.
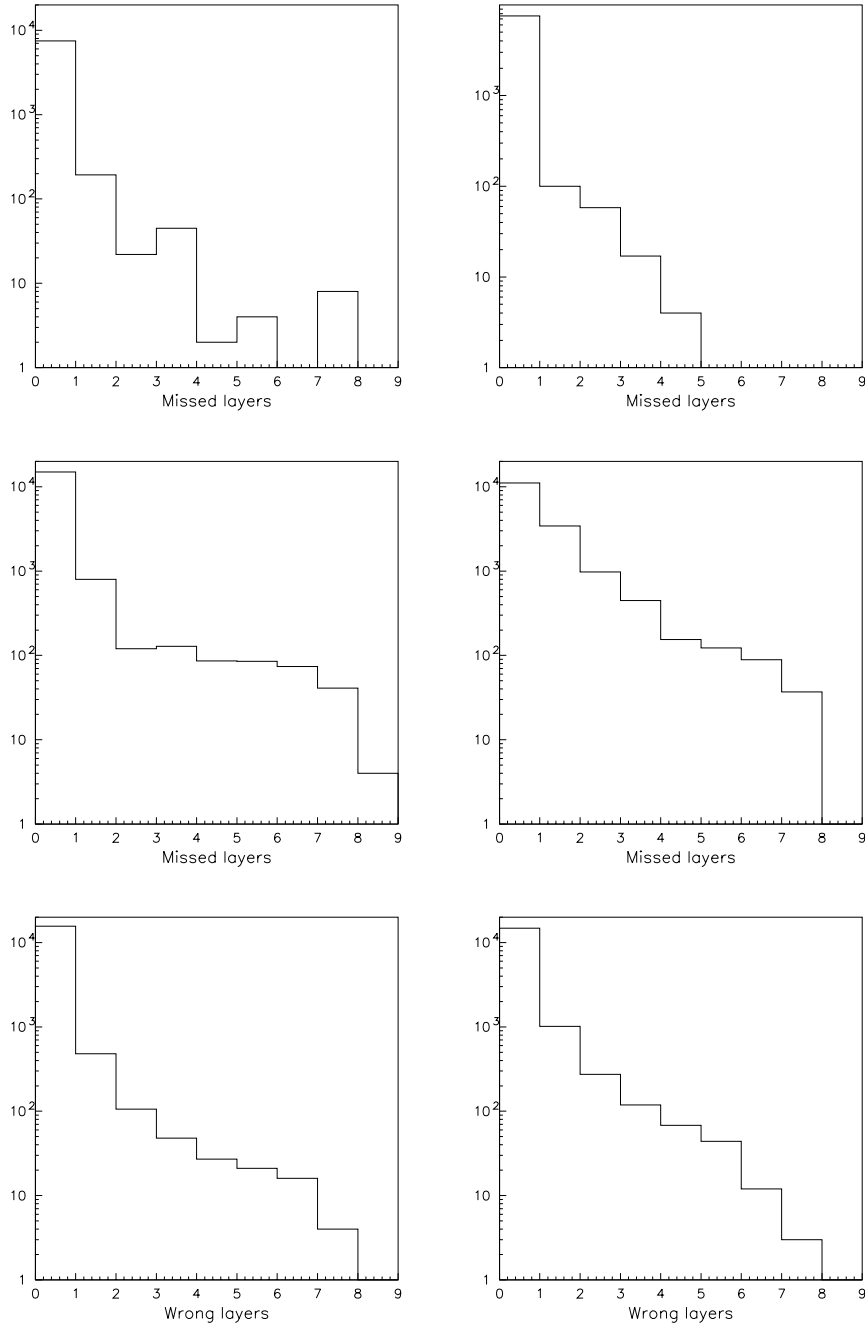
Figure 8.9:   Missed axial layers in silicon reconstruction on single track (above) and $t\bar{t}$ (center) events; wrong layers in $t\bar{t}$ events (below). Performance of Outside-In Tracking (left) is compared to Histogram Tracking (right).
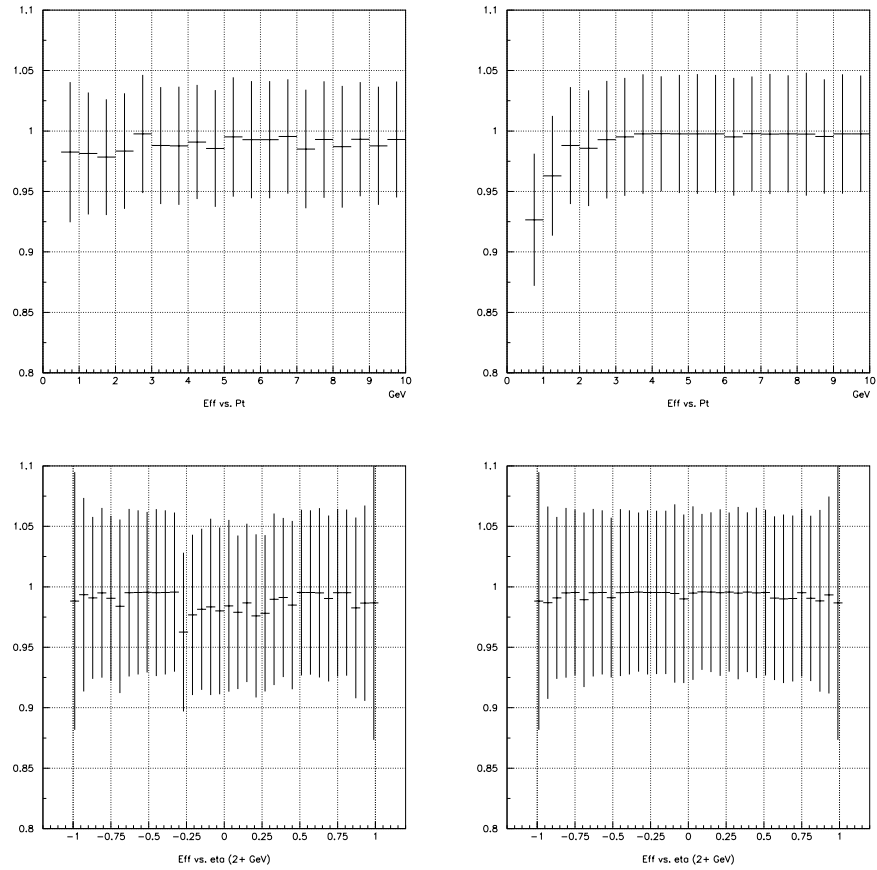
Figure 8.10: Fit-based efficiency on single-track events, as a function of track parameters, in Outside-In (left) and Histogram (right) axial reconstruction. Efficiency vs η is computed for the tracks with p_t above 2 GeV.
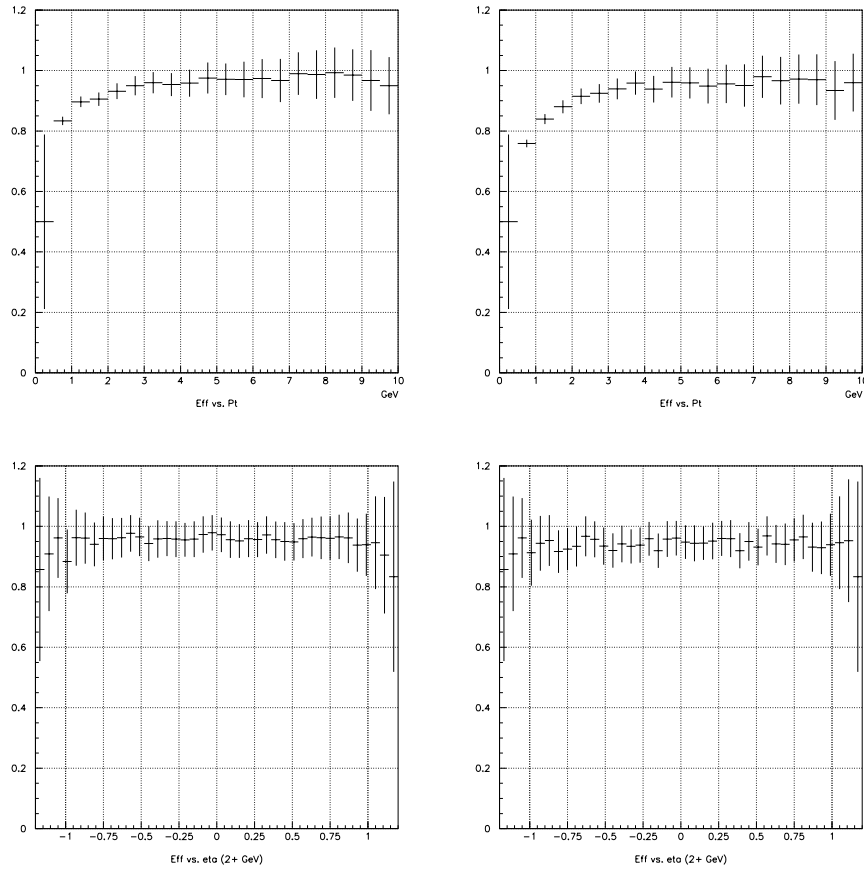
Figure 8.11: Fit-based efficiency on $t\bar{t}$ events, as a function of track parameters, in Outside-In (left) and Histogram (right) axial reconstruction. Efficiency vs η is computed for the tracks with p_t above 2 GeV.
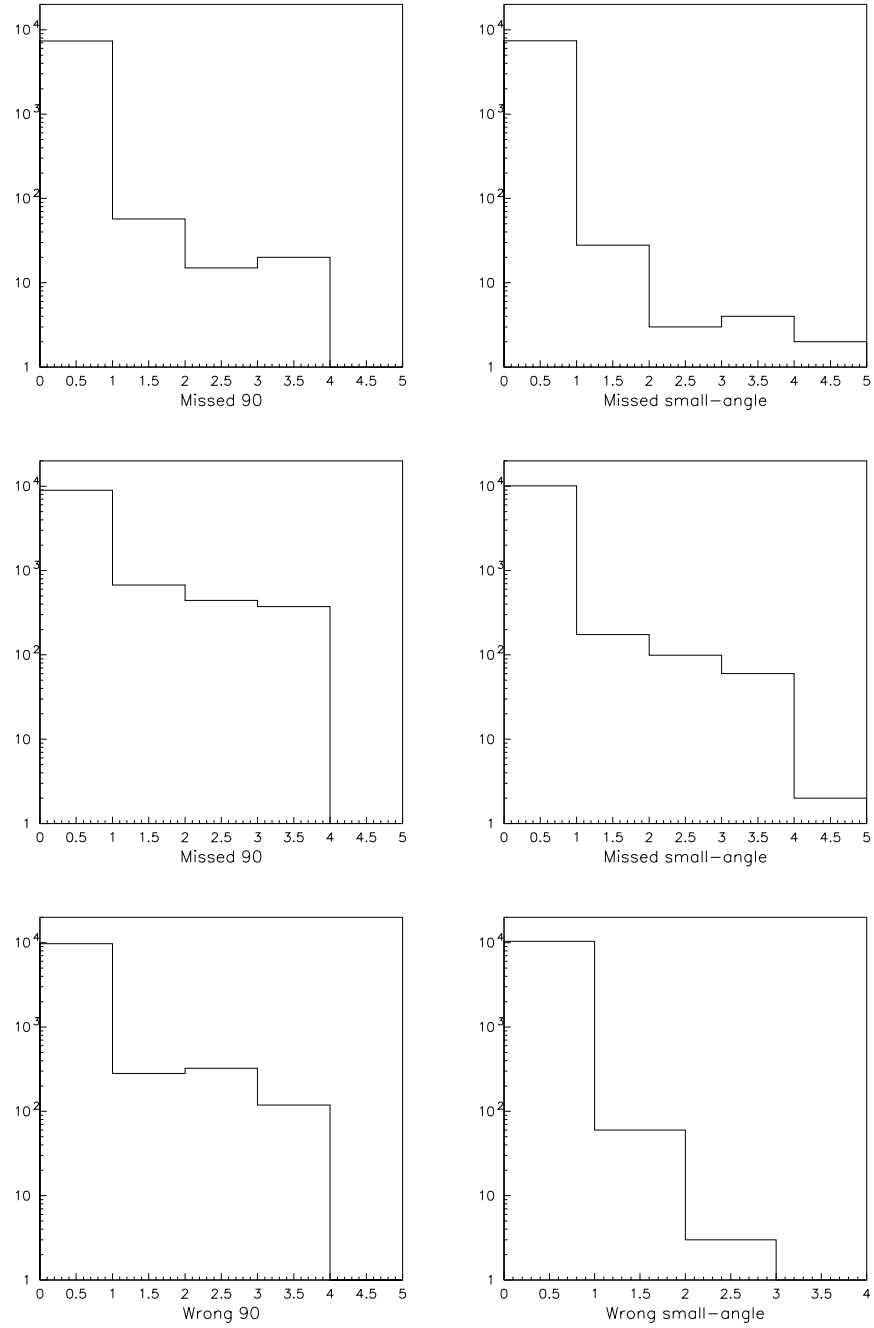
Figure 8.12: Missed stereo layers in silicon reconstruction on single track (above) and $t\bar{t}$ (center) events; wrong layers in $t\bar{t}$ events (below). 90° and small-angle layers are reported separately. Outside-In (left) is compared to Histogram Tracking (right).
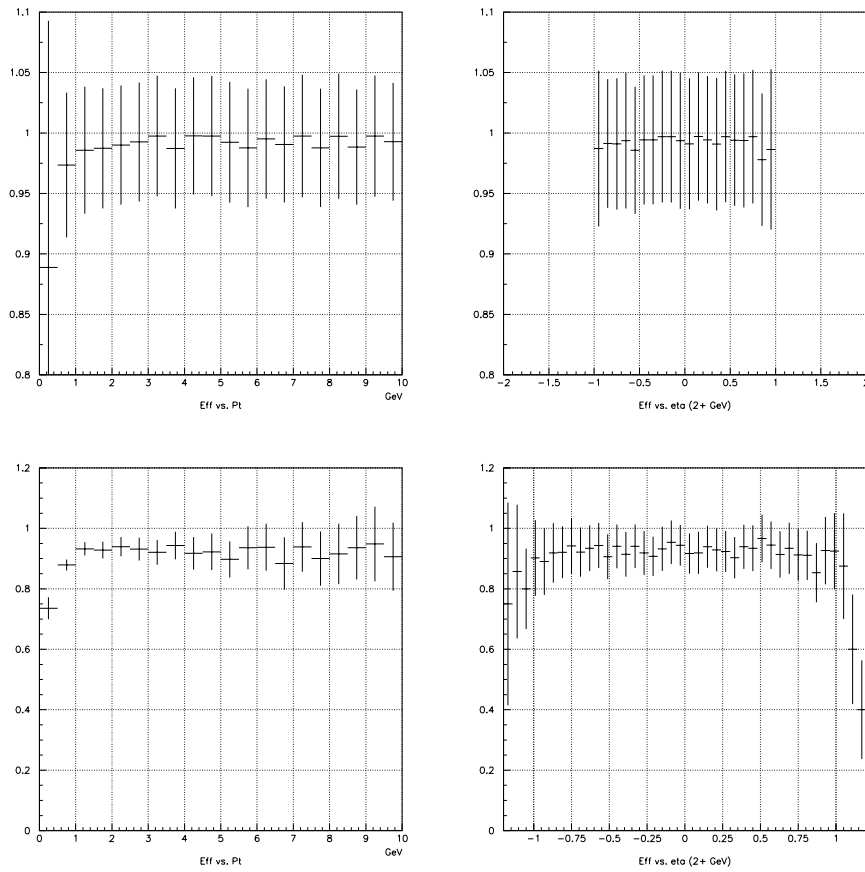
Figure 8.13:   Fit-based efficiency on single-track (above) and $t\bar{t}$ (below) events, as a function of track parameters, in Histogram stereo reconstruction. Efficiency vs η is computed for the tracks with p_t above 2 GeV.

Conclusions

This thesis is a study on the various aspects of track reconstruction at the CDF II experiment, ranging from fast tracking at trigger level to precision tracking in offline analyses.

Part of my work was devoted to studying SVT, the silicon track processor which is part of the second level of trigger at CDF II. The standalone, bit-level simulation of SVT was interfaced to the complete simulation of the detector, and used to predict SVT's performance.

At first, the resolution, efficiency and fake rate of the tracker were estimated using the experiment's baseline configuration. These results were then compared to the performance of two other configurations, which make use of the recently-added "Layer 00" — a radiation-hard silicon layer placed immediately outside the beam pipe. As expected, introduction of Layer 00 improved the track resolution considerably with respect to the baseline; one of the two configurations was also shown to decrease the fake rate.

At the third level of trigger, and in offline analyses, CDF II reconstructs tracks by means of several strategies, which are part of a C++ framework. My main research subject was the development and implementation of a general-purpose tracking strategy, called "Histogram Tracking", which was then applied both to the drift chamber and to the silicon vertex detector. This strategy is now part of the official tracking code of CDF II; it will be used to reconstruct real events when the experiment will begin to collect data, in spring 2001.

In order to evaluate the performance of Histogram Tracking and of other strategies, I developed a series of benchmarks, based on "perfect track reconstruction". I implemented an ideal pattern finder, which makes use of parentage data from event simulations; I used this tool to determine a lower bound on the resolution of track parameters, as a function of momentum and detector occupancy. I defined two complementary sets of benchmarks: one compares the hit content of found tracks to the corresponding ideal tracks, while the other relies on the accuracy of fit parameters.

The final step consisted in applying the benchmarks to CDF II's tracking

strategies, and in estimating the cost of each strategy in terms of CPU time.

For the purpose of track reconstruction in the transverse plane of the drift chamber, Histogram Tracking turned out to be a very reliable algorithm, with a high efficiency even in high-occupancy situations. The complementary strategy, Segment Linking, while being faster, is more sensitive to occupancy.

Once a track has been found in the drift chamber, it is extrapolated towards the silicon vertex detector, and it is associated to a set of silicon hits. In this situation, Outside-In — a strategy based on Kalman filtering — proved to be the most accurate strategy; however, it is also very slow. Histogram Tracking requires less time; it is less reliable in borderline decisions, but most of the Histogram Tracking errors have a small impact on the fit parameters of reconstructed tracks. Therefore, silicon Histogram Tracking is being considered for use at trigger level, where speed is essential.

CDF II will start to collect data in a few months. Very exciting results are expected to come soon, thanks to the detector upgrade, and in particular to the new tracker. For example, by making use of the track reconstruction techniques that were studied in this thesis, CDF II should be able to measure B_s oscillations by the end of 2001, for values of x_s up to about 30.

Bibliography

[1] Observation of a Dimuon Resonance at 9.5 GeV in 400-GeV Proton-Nucleus Collisions
S. W. Herb et al.
Phys. Rev. Lett. **39**, 252 (1977)

[2] B Lifetimes, Mixing and CP Violation at CDF
M. Paulini
Int. J. Mod. Phys. **A14**, 2791 (1999)

[3] Summary of CDF b-quark integral cross-section measurement
CDF Collaboration (1997)
`http://www-cdf.fnal.gov/physics/new/bottom/cdf1991_new/`
`xsec_int_color_1997.ps`

[4] Production of three large-p_t jets in hadron-hadron collisions
Z. Kunszt, E. Pietarinen
Nucl. Phys. **B164**, 45 (1980)

[5] Improved parton distributions from global analysis of deep inelastic scattering and inclusive jet data
H. L. Lai et al. (CTEQ collaboration)
Phys. Rev. **D55**, 1280 (1997) **hep-ph/9606399**

[6] Measurement of the proton and the deuteron structure functions F_2^p and F_2^d
M. Arneodo et al. (NMC collaboration)
Phys. Lett. **B364**, 107 (1995)

[7] A measurement of the proton structure function $F_2(x, Q^2)$
T. Ahmed et al (H1 collaboration)
Nucl. Phys. **B439**, 471 (1995)

[8] Measurement of the proton structure function F_2 from the 1993 HERA data
M. Derrick et al. (ZEUS collaboration)
Z. Phys. **C65**, 379 (1995)

[9] Inclusive Jet Cross Section in $p\bar{p}$ Collisions at $\sqrt{s} = 1.8$ TeV
F. Abe et al. (CDF Collaboration)
Phys. Rev. Lett. **77**, 438 (1996)

[10] Status of fragmentation models
T. Siöstrand
Int. J. Mod. Phys **A3**, 751 (1988)

[11] Scaling violations in inclusive e^+e^- annihilation spectra
C. Peterson et al.
Phys. Rev. **D27**, 105 (1983)

[12] Search for $B \to \ell\bar{\nu}_\ell$
Gao et al. (CLEO Collaboration)
Phys. Rev. Lett. **75**, 785 (1995)

[13] Lifetime of Particles Containing b quarks
E. Fernandez et al. (MAC collaboration)
Phys. Rev. Lett. **51**, 1022 (1983)

[14] Measurement of the Lifetime of Bottom Hadrons
N. Lockyer et al. (MARKII collaboration)
Phys. Rev. Lett. **51**, 1316 (1983)

[15] LEP B lifetimes working group
http://claires.home.cern.ch/claires/lepblife_page2.html

[16] Observation of B_c Mesons in $p\bar{p}$ Collisions at $\sqrt{s} = 1.8$ TeV
F. Abe et al. (CDF Collaboration)
Phys. Rev. **D58**, R112004 (1998)

[17] A precise measurement of the average b hadron lifetime
D. Buskulic et al. (ALEPH Collaboration)
Phys. Lett. **B369**, 151 (1996)

[18] Measurement of the average lifetime of b-hadrons in Z decays
M. Acciarri et al. (L3 Collaboration)
Phys. Lett. **B416**, 220 (1998)

[19] Measurement of the average b hadron lifetime in Z^0 decays
P. D. Acton et al. (OPAL Collaboration)
Zeit. Phys. **C60**, 217 (1993)

[20] Measurement of the b hadron lifetime with the dipole method
D. Buskulic et al. (ALEPH Collaboration)
Phys. Lett. **B314**, 459 (1993)

[21] A precision measurement of the average lifetime of B hadrons
P. Abreu et al. (DELPHI Collaboration)
Zeit. Phys. **C63**, 3 (1994)

[22] Updated precision measurement of the average lifetime of B hadrons
P. Abreu et al. (DELPHI Collaboration)
Phys. Lett. **B377**, 195 (1996)

[23] A measurement of the average bottom hadron lifetime
K. Ackerstaff et al. (OPAL Collaboration)
Zeit. Phys. **C73**, 397 (1997)

[24] Measurement of the Average B Hadron Lifetime in Z^0 Decays Using Reconstructed Vertices
K. Abe et al. (SLD Collaboration)
Phys. Rev. Lett. **75**, 3624 (1995)

[25] Measurement of B hadron lifetimes using J/ψ final states at CDF
F. Abe et al. (CDF Collaboration)
Phys. Rev. **D57**, 5382 (1998)

[26] Behaviour of Neutral Particles under Charge Conjugation
M. Gell-Mann and A. Pais
Phys. Rev. **97**, 1387 (1955)

[27] Measurement of the $B^0\bar{B}^0$ Oscillation Frequency Using $\pi - B$ Meson Charge-Flavor Correlations in $p\bar{b}$ Collisions at $\sqrt{s} = 1.8$ TeV
F. Abe et al. (CDF Collaboration)
Phys. Rev. Lett. **80**, 2057 (1998)

[28] Measurement of the $B_d^0\bar{B}_d^0$ Flavor Oscillation Frequency and Study of Same Side Flavor Tagging of B Mesons in $p\bar{p}$ Collisions
F. Abe et al. (CDF Collaboration)
Phys. Rev. **D58**, 032001 (1999)

[29] Measurement of $B^0\bar{B}^0$ Flavor Oscillations Using Jet-Charge and Lepton Flavor Tagging in $p\bar{p}$ Collisions at $\sqrt{s} = 1.8$ TeV
F. Abe et al. (CDF Collaboration)
Phys. Rev. **D60**, 072003 (1999)

[30] Measurement of the $B_d^0\bar{B}_d^0$ Oscillation Frequency Using Dimuon Data in $p\bar{p}$ Collisions at $\sqrt{s} = 1.8$ TeV
F. Abe et al. (CDF Collaboration)
Phys. Rev. **D60**, 051101 (1999)

[31] Measurement of the $B_0\bar{B}_0$ Oscillation Frequency using $\ell^- D^{*+}$ Pairs and Lepton Flavor Tags
T. Affolder et al. (CDF Collaboration)
Phys. Rev. **D60**, 112004 (1999)

[32] B Oscillation Results for the BEAUTY 2000 conference
LEP Working Group on B Oscillations
http://lepbosc.web.cern.ch/LEPBOSC/combined_results/sept_2000/

[33] Evidence for the 2π decay of the K_2^0 meson
J. H. Christenson, J. W. Cronin, V. L. Fitch and R. Turlay
Phys. Rev. Lett **13**, 138 (1964)

[34] CP Violation in the Renormalizable Theory of Weak Interactions
M. Kobayashi, T. Maskawa
Progr. Theor. Phys **49**, 652 (1973)

[35] Electroweak Symmetry Breaking and Higgs Physics
M. Spira, P.M.Zerwas
hep-ph/9803257 (1997)

[36] Parametrization of the Kobayashi–Maskawa Matrix
L. Wolfenstein
Phys. Rev. Lett. **51**, 1945 (1983)

[37] A Measurement of $\sin 2\beta$ from $B \to J/\psi K_s^0$ with the CDF Detector
T. Affolder et al. (CDF Collaboration)
Submitted to Phys. Rev. D September 7, 1999

[38] First CP Violation Results from BaBar
D. Hitlin (BaBar Collaboration)
Presented at ICHEP 2000, Osaka
hep-ex/0011024

[39] Quark mixing, CP violation and rare decays after the top quark discovery
A. Buras, R. Fleischer
`hep-ph/9704376`

[40] Observation of $B \to K^{\pm}\pi^0$ and $B \to K^0\pi^0$, and Evidence for $B \to \pi^+\pi^-$
D. Cronin-Hennessy et al. (CLEO Collaboration)
Phys. Rev. Lett. **85**, 515 (2000)

[41] The CDF detector: an overview
F. Abe et al. (CDF Collaboration)
FERMILAB-PUB-88-25-E, Feb 1988. 40pp.
Nucl. Instrum. Meth. **A271**, 387-403 (1988)

[42] List of CDF I physics publications and preprints:
`http://www-cdf.fnal.gov/physics/preprints/index.html`

[43] Evidence for Top Production in $p\bar{p}$ Collisions at $\sqrt{s} = 1.8$ TeV
F. Abe et al. (CDF Collaboration)
Phys. Rev. Lett. **73**, 225 (1994)

[44] Evidence for Top Quark Production in $p\bar{b}$ Collisions at $\sqrt{s} = 1.8$ TeV
F. Abe et al. (CDF Collaboration)
Phys. Rev. **D50**, 2966 (1994)

[45] Kinematic evidence for top quark pair production in W+multijet events in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV
F. Abe et al. (CDF Collaboration)
Phys. Rev. **D51**, 4623 (1995)

[46] Measurement of the Top Quark Mass with the Collider Detector at Fermilab
T. Affolder et al. (CDF Collaboration)
FERMILAB-PUB-00/127-E
Submitted to Phys. Rev. D June 29, 2000

[47] Measurement of the W Boson Mass with the Collider Detector at Fermilab
T. Affolder et al. (CDF Collaboration)
Submitted to Phys. Rev. D July 26, 2000

[48] RUN II HANDBOOK
`http://www-bd.fnal.gov/lug/runII_handbook/RunII_index.html`

[49] CDF II Technical Design Report
     CDF Collaboration
     FERMILAB-PUB-96/390-E, Nov 1996. 318pp.
     http://www-cdf.fnal.gov/upgrades/tdr/tdr.html

[50] Design and construction of the CDF central tracking chamber
     F. Bedeschi et al.
     FERMILAB-PUB-87/182-E, Jul 1987. 80pp.
     Nucl. Instrum. Meth. **A268**, 50 (1988)

[51] The CDF central electromagnetic calorimeter
     L. Balka et al.
     FERMILAB-PUB-87-172-E, (Received Dec 1987). 26pp.
     Nucl. Instrum. Meth. **A267**, 272 (1988)

[52] The CDF central and endwall hadron calorimeter
     S. Bertolucci et al.
     FERMILAB-PUB-87-174-E, Jul 1987. 32pp.
     Nucl. Instrum. Meth. **A267**, 301 (1988)

[53] Shower maximum trigger for electrons and photons at CDF
     K.Byrum et al.
     NIMPR **A364**, 144 (1995)

[54] Precision Measurement of the Prompt Photon Cross Section in $p\bar{b}$ Collisions at $\sqrt{s} = 1.8$ TeV
     F. Abe et al. (CDF Collaboration)
     Phys. Rev. Lett. **73**, 2662 (1994)

[55] The CDF forward muon system
     K. Byrum et al.
     FERMILAB-PUB-87/181-E, Jul 1987. 16pp.
     Nucl. Instrum. Meth. **A268**, 46 (1988)

[56] CDF central muon detector
     G. Ascoli et al.
     FERMILAB-PUB-87/179-E, Jul 1987. 33pp.
     Nucl. Instrum. Meth. **A268**, 33 (1988)

[57] CDF Run II Code Browser
     http://purdue-cdf.fnal.gov/CdfCode/

[58] Tracking in the SVX
Hans Wenzel, Fermilab
CDF internal note CDF/DOC/SEC_VTX/PUBLIC/1790 (1998)

[59] The CDF Run II Silicon Tracking System
A. Sill et al.
Nucl. Instrum. Meth. **A447**, 1 (2000)

[60] Simulation of the XFT II Trigger
R. Hughes, P. Koehn, B. Winer, G. Veramendi
`http://www.physics.ohio-state.edu/~hughes/cdf_osu/xft/`
`talks/standalone_sim.ps`

[61] XFT II Status
B. Winer et al.
`http://www.physics.ohio-state.edu/~hughes/cdf_osu/xft/`
`talks/xftii_status.ps`

[62] Silicon Vertex Tracker: Technical Design Report
S. Belforte et al.
`http://www.pi.infn.it/cdf/svt/svt_tdr.ps`

[63] Performances comparison of SVT in the baseline configuration and with
Layer 00
G. Busetto, D. Lucchesi, G. Punzi
CDF internal note CDF/DOC/TRIGGER/CDFR/5134 (1999)

[64] Review of Particle Physics: Statistics
Particle Data Group
The European Physical Journal C (1988), 172—177

[65] Application of Kalman Filtering to Track and Vertex Fitting
R. Frühwirth
Nuclear Instruments **A262**, 444 (1987)

List of Figures

List of Tables

Index

Acknowledgements

This thesis, and the underlying research, would not have been completed without the contribution of many friends and colleagues.

First and foremost I must thank my official and non-official advisors, Dario Bisello and Giovanni Busetto. They gave me a warm welcome into the Padova CDF group; their patience, their trust and their good counsel have meant much to me.

Donatella Lucchesi and Irene Fiori shared my work on the Silicon Vertex Tracker. Donatella gave me several interesting suggestions regarding the physics involved in SVT, while Irene spent plenty of time researching the best way to generate events with the correct amount of noise. Both of them were quite helpful.

SVT itself would not exist without the SVT group in Pisa. I'd like to thank Giovanni Punzi, Luciano Ristori and Simone Donati for their suggestions and the information they provided. Guram Chlachidze maintained and extended Giovanni's SVT simulator, and provided us with the sets of associative memory patterns we needed.

When he was an Erasmus student in Padova, David Wahl worked on the preliminar SVT testing procedure. He was of great help in setting up the simulation, and located several bugs and quirks. Unfortunately for him, the code began to give trustworthy results only half an hour after he presented his final report!

I am quite grateful to Avi Yagil for the many discussions we had about Histogram Tracking, and for suggesting I should investigate this approach in the first place. Avi also asked me to develop several "sanity check" tools, which later on proved quite useful in ensuring we were obtaining meaningful results — or understanding why, when we didn't.

Alberto Ribon was the first member of the Padova CDF group to study Histogram Tracking. The current code is an extension and generalization of his silicon rz reconstruction. Even if he does not work with us any longer, Alberto is unforgettable.